DOCUMENT RESUME

ED 459 861 IR 058 413

| | |
|---|---|
| AUTHOR | Monaco, Michael J. |
| TITLE | A Model Privacy Statement for Ohio Library Web Sites. |
| PUB DATE | 2001-05-00 |
| NOTE | 70p.; Master of Library and Information Science Research Paper, Kent State University. |
| PUB TYPE | Dissertations/Theses (040) |
| EDRS PRICE | MF01/PC03 Plus Postage. |
| DESCRIPTORS | *Electronic Libraries; Information Seeking; *Library Policy; *Library Services; Models; *Privacy; *Standards; Users (Information); *World Wide Web |
| IDENTIFIERS | American Library Association; Federal Trade Commission; HTML; Organisation for Economic Cooperation Development; Web Site Design; *Web Sites |

ABSTRACT

The purpose of this research was to develop a model privacy policy statement for library World Wide Web sites. First, standards of privacy protection were identified. These standards were culled from the privacy and confidentiality policies of the American Library Association, the Federal Trade Commission's online privacy reports, the guidelines of the Organization for Economic Cooperation and Development, and the writings of particular information professionals. Then, these standards were used to identify the privacy issues raised by various information seeking processes, e.g., cookies, clickstream and log files, e-mail, interlibrary loans, and searches. With the issues identified, the standards of privacy protection were used to make recommendations for libraries to ensure that the privacy of library users is adequately protected. The recommendations were used to construct a model privacy statement for library Web sites. Appendices include the model policy, the HTML source of the model policy, and the HTML source of the policy annotation. (Contains 18 references.) (Author/MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 459 861

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
D. P. Wallace
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)



# A MODEL PRIVACY STATEMENT FOR OHIO LIBRARY WEB SITES

A Master's Research Paper submitted to the
Kent State University School of Library
and Information Science
in partial fulfillment of the requirements
for the degree Master of Library and Information Science

by

Michael J. Monaco

May, 2001

IR058413

BEST COPY AVAILABLE

Master's Research Paper by

Michael J. Monaco

B.A., University of Toledo, 1994

M.A., Kent State University, 1996

M.L.I.S, Kent State University, 2001

Approved by

Advisor Thomas J. Frosklak Date May 3, 2001

## CONTENTS

PREFACE ..... iv

Chapter

I. INTRODUCTION ..... 1

II. GENERAL STATEMENTS OF POLICY AND THE SCOPE OF CONFIDENTIALITY/PRIVACY ..... 2

Privacy and Confidentiality Policies of the ALA

The FTC's Online Privacy Reports

Guidelines of the Organization for Economic Cooperation and Development

Personal Statements of Privacy Policies

III. SPECIFIC PRIVACY ISSUES OF THE ONLINE ENVIRONMENT ..... 17

Cookies

Clickstream and Log Files

Email

Interlibrary Loans

Searches

IV. CONCLUSIONS ..... 30

NOTES ..... 32

WORKS CITED ..... 34

APPENDIX ONE : THE MODEL POLICY ..... 36

APPENDIX TWO : HTML SOURCE OF THE MODEL POLICY ..... 41

APPENDIX THREE : HTML SOURCE OF THE POLICY ANNOTATION ..... 47

## PREFACE

This paper, submitted as a physical object, does not reflect the entirety of the work that this project includes. The three appendices give, as accurately as possible, a picture of the computer files my work culminated in. The first appendix is a printout of the HTML model policy. It is included to give an idea of what the policy looks like on a computer screen. This policy has links to several internet sites and to passages in an HTML version of the present paper. The second appendix presents the HTML source that comprises the model policy, thus revealing the links and their targets. The third appendix is the HTML source for the paper, presented in order to reveal the passages which are intended to be linked to the policy as annotations explaining the rationale for specific policy provisions and to guide librarians who may customize the model policy to create a policy statement for their library's web sites.

I would like to thank several people, without whose assistance and support I would not have been able to complete my program of study or the present paper. I thank Thomas Froehlich, my research advisor, for helping me focus and find a manageable topic, and also for setting a fine model of how a philosopher can raise difficult but necessary questions in information science. I thank Karen Coyle for generously offering her own work on this topic for my edification. I thank all of my instructors in the School of Library and Information Science at Kent State University for the knowledge and love of the discipline they have imparted to me.

But any professional work is possible only with the help of others who may have no special interest or expertise in the profession, yet provide support and inspiration on a personal

level. I thank Deborah for her patience, her love, and for believing in me. I thank my parents for their enduring support and enthusiasm for my studies. I thank my brothers Tom and Joe, and my sisters Lisa and Lynn, for their humor and sympathetic ears.

# I. INTRODUCTION

"Privacy" is a complex concept which has taken many meanings over time. In the past, privacy was understood to be a right to freedom from observation and scrutiny in the personal life of an individual, so that privacy was felt to be violated when a person's behavior was observed or otherwise made public in a way the person found invasive and revealing. Thus, an individual's life within the home and among intimate acquaintances was understood to be protected to some extent from public scrutiny. But nowadays privacy is also thought to extend to information about a person, especially information which related to behaviors or status that is not regarded as public. Thus financial records, health records, and the like are thought to be "private" in the sense that such information is or should be under the control of the subject of this information. It is also commonly held that intimate, personal exchanges between an individual and various professionals who are given personal information about the individual should be regarded as confidential. Because information seeking behaviors can potentially reveal private information about the seeker, and because such information seeking often is assisted by information professionals with an understanding of confidentiality, librarians and other information professionals must recognize and protect the confidentiality of any records of information seeking. This paper will explore guidelines outlining the protection of privacy and specific kinds of information that are generated in information seeking, in order to develop a set of recommendations for the protection of the confidentiality of library users in the online environment.

## II. GENERAL STATEMENTS OF POLICY AND THE SCOPE OF CONFIDENTIALITY/PRIVACY

Confidentiality and privacy are values which have been linked to both freedom of expression and the right to access information. Librarians and information professionals in general recognize the importance of maintaining the trust of information seekers, and have undertaken a variety of efforts to insure that the confidence and privacy of their clients are protected. As the internet and online environments have become increasingly important sources of information, they have also created a number of new contexts for privacy issues. This paper will identify and explore the privacy issues which accompany online information seeking, and what steps librarians should take to address these issues. To see that privacy is of importance to librarians, this paper will exhibit the policies of the ALA and some statements made by individual librarians and information professionals. Then, what privacy protection consists in will be identified, and finally these principles will be applied to specific issues in the online environment.

### Privacy and Confidentiality Policies of the ALA

The American Library Association's policy manual, section 54.16 "On Professional Ethics," offers a series of statements to guide ethical conduct for librarians, one of which is "We protect each library user's right to privacy and confidentiality with respect to information sought or

received and resources consulted, borrowed, acquired, or transmitted."[1] This simple statement establishes that a core value of librarianship is the protection of every user's privacy. The library user may not think very much about confidentiality when seeking information, but even so there is a duty to safeguard the user's privacy and confidentiality. It is not this paper's purpose to explore the reasons and justifications supporting this duty on the part of libraries, although it is clear that in part this duty arises from the potential "chilling effect" a lack of protection might create among users. The important point is that the ALA recognizes confidentiality and privacy as important rights of library users.

One area where the confidentiality of library users has been challenged is in circulation records. Law enforcement agencies have requested this information, either to obtain the records of specific persons or as part of "fishing expeditions" to identify potential suspects. The ALA has repeatedly fought such efforts and section 52.4 specifically addresses the confidentiality of library records. In fact, 52.4 identifies a variety of records, beyond just circulation records, for protection:

> Confidentiality extends to "information sought or received, and materials consulted, borrowed, or acquired," and includes database search records, reference interviews, circulation records, interlibrary loan records, and other personally identifiable uses of library materials, facilities, or services.[2]

There are several important ideas in this sentence. First, it states that confidentiality extends to any library record that could identify particular persons. Second, it states that confidentiality extends to uses of library resources, suggesting that any use of resources which produces some type of record is to be protected, not just the uses of resources that are mediated by library personnel. Thus, if a library user were to use a library computer to seek information, any

electronic records of such use (and we shall see that there are several) should also be protected. Finally, it should be noted that there is an emphasis on personally identifiable information -- that is, information which can be linked with specific individuals. Where records of information seeking are aggregated, abstracted, or otherwise collected into data that does not identify specific individuals in a way that can be traced back to individuals, this information can rightly be regarded as falling outside the concerns for and protections of privacy and confidentiality.

Section 52.4 of the policy manual ("Confidentiality of Patron Records") addresses the appropriate legal procedure for law enforcement to access such confidential records, but it also stresses the need for libraries to adopt formal policies recognizing the above standard of confidentiality. In order to adequately address the confidentiality of records, libraries must address the specific kinds of records created in the online environment, which will be discussed below.

The ALA's Intellectual Freedom Manual reiterates this policy under the headings "Policy on the Confidentiality of Library Records" and "Policy concerning Confidentiality of Personally Identifiable Information about Library Users."[3] These two sections actually present portions of 52.4, and also provide some historical background on the events which led the ALA to issue the policy. Apart from the historical accounts, the Intellectual Freedom Manual does not alter or expand the scope or range of records identified in the ALA Policy Manual.

Apart from the general policies of the ALA, there is another document outlining the impact of computer technology on the privacy and confidentiality of patron records. This document is the Report of the Task Force on Privacy and Confidentiality in the Electronic

Environment, issued July 7, 2000.[4] The report includes a number of important findings.

First, there is the issue of computer system security. The Task Force found that insofar as some patron records exist as electronic records on library computers, the library must recognize a duty to protect these records from hackers or other unauthorized access. Specific security measures fall outside the scope of my work but some account of the library's commitment to resist unauthorized access to computer records should be disclosed to users.

Second, the Task Force found that a variety of privacy issues are raised by in-library internet access. Every internet transaction is potentially a source of information gathering by web sites and servers outside the library. Because many patrons may use a computer over the course of a day, there needs to be some awareness of the potential for subsequent users to see traces of the web sites visited by previous users. There is also the potential for other patrons to see the screen of a computer being used by a patron, thereby compromising the user's privacy.

Third, the Task Force recognizes potential problems raised by the use of electronic resources provided by a remote vendor. These vendors may collect data about resource access and even about the patrons accessing them without guaranteeing the same protections of a library. Thus, it is important for libraries to establish privacy protection measures in their licenses and contracts with vendors or else to give patrons notice that personal information may be collected by the vendor.

Apart from these significant findings, the Task Force issued a number of recommendations. Beyond revising ALA policies concerning confidentiality of records to address the findings mentioned above, the Task Force calls on the ALA to "urge that all libraries

adopt a privacy statement on web pages and post privacy policies in the library which cover the issues of privacy in internet use as accessed through the library's services."[5]

**The FTC's Online Privacy Reports**

The Federal Trade Commission's reports on developing a theoretical framework for fair information practices online stand out as particularly explicit and well-informed documents. The focus of the FTC's two reports to Congress has been on commercial use of the internet, but the general discussion of fair information practices here is particularly useful. The FTC reviewed some twenty-five years worth of government documents and studies concerning the collection of personal information and identified five core principles of privacy which were common to all.[6] These principles were identified in 1998 as Notice/Awareness, Choice/Consent, Access/Participation, Integrity/Security, and Enforcement/Redress. Before discussing the particulars of each principle, it should be noted that the FTC's 2000 report to Congress pares down the list Notice, Choice, Access, and Security -- Enforcement/Redress is dropped from the core[7] -- presumably because the 2000 report is intended to guide enforceable legislation, and not because the FTC no longer views the enforcement of privacy protection as important.

The principle of notice is regarded as the most fundamental of the principles.[8] The idea is that the individual must have an awareness of the practices of information collecting agencies before he can decide to disclose personal information. In addition, the individual would need to know that information is collected before he could attempt to access, evaluate, or correct the information, and also before he could make any effort to redress illicit or unauthorized use of the

information. The FTC concludes that the principle of notice requires that personal information collecting entities should be required to disclose at least some of the following in order to provide meaningful notice: (1) the identity of the entity collecting data; (2) the uses to which data is put; (3) the identity of any potential recipients of the data; (4) the nature of the data collected and means of collection; (5) whether the individual user or consumer is required to furnish personal information, and the consequences of refusal to do so; and (6) the measures taken by the collecting entity to protect the data, ensure its accuracy, and uphold its confidentiality.[9] The FTC's requirements are not quite stringent enough for libraries, in so far as libraries have a positive commitment to protecting confidentiality, and so all of these items should be disclosed in a library privacy policy. The FTC does however raise an important point about the posting of policies: they should be readily assessable both from the home page of a web site and from any pages where personal information is collected.[10]

The principle of choice requires that the individual whose personal information is to be collected be given the option to furnish information or not. It also requires that the individual be given a choice about how the information will be used, and especially about how the information may be put to use beyond the initial transaction that requires the disclosure of personal information.[11] Thus, if the owner of a web site intends to collect personal information, there must be consent from the user or consumer both for the initial collection of information and for any uses the information might be put to: web site management and development, sale to advertisers or direct marketing firms, sharing with other companies for market research, or any other use. This principle can be implemented in one of two ways -- either allowing the user to opt-in or to opt-out. "Opt-in" means the individual must actively consent to the collection and

use of information. "Opt-out" means that the user must actively deny consent for information collection and use. It is hardly surprising that most collectors of information prefer the latter, as it allows them to assume consent from any individual, and requires special effort on the individual's part, rather than on the information collecting entity's part, to protect privacy. Libraries, however, would probably show greater fidelity to the user's confidentiality by using "opt-in" protocols wherever possible, as this provides an additional opportunity to both disclose policy and assure the user of the library's commitment to privacy.

The principle of access is comprised of two rights on the part of the individual about whom information is collected: the right to review any information and the right to correct or contest the information.[12] This principle can only be upheld effectively if there is a relatively easy and inexpensive procedure for reviewing and correcting the data. The more this information is limited in scope and content, the easier it will be to ensure that the information is accurate.

The principle of security covers both the integrity or quality of information and the protection of information from loss and the unauthorized disclosure, use, and access of the information.[13] The integrity of information is protected in part by allowing the individual to access and correct information, but there is also a responsibility on the part of the collecting entity to use prudence in selecting sources of information. Finally, the occasional destruction of obsolete or old information is necessary to protect the quality and integrity of information. Protecting the security of information calls for both technological and administrative measures.[14] The technological measures involved might include encryption, pass-word protection, firewalls, and storage on secure servers. The administrative measures should focus on setting limits on

who can access the information and to what purposes the information can be used. Obviously, it is imperative that the information never be used in manners or for purposes not authorized by the individual. Libraries should have policies in place which limit access to patron records of any kind, and also have technological protections sufficient to prevent unauthorized access to records.

The last principle is the principle of enforcement. The FTC concludes that "the core principles of privacy can only be effective if there is a mechanism in place to enforce them."[15] The FTC outlines several mechanisms of enforcement: self-regulation, private remedies, and government enforcement. Self-regulation, the FTC contends, can be carried out by requiring members of professional or industrial organizations to accept and comply with a privacy policy. There is a clear application of this principle for libraries: the adoption of privacy policies by organizations such as the ACRL, the ALA, and the SLA. These policies should be fairly explicit, and there would also be a need for developing a system to verify compliance. Because it is unlikely that libraries would be motivated to violate privacy policies (unlike, say, private firms which might realize a profit from selling customer information), it may be enough to simply require libraries to adopt formal privacy policies and post them on their web sites. Tort law provides for a certain amount of private remedies. But it is unclear, at this point, what steps legislatures may take to pass laws concerning privacy protection.

Taken together, the five principles of privacy protection identified by the FTC provide a strong foundation for specifying the principles libraries should uphold with respect to personal information. But it is also helpful to notice to some specific statements of principles offered by librarians and other information professionals concerning privacy and confidentiality. This is

important to do because the FTC's principles constitute a minimal set of guidelines to restrict the activities of commercial entities. But librarians are committed to doing more than merely setting limits on their use of personal information -- they must take positive steps to ensure that confidentiality is always protected.

**Guidelines of the Organization for Economic Cooperation and Development**

Although it is impossible and unnecessary to discuss all of the many statements of principles of privacy protection created by organizations and individuals, it is enlightening to review the guidelines issued in 1980 by the OECD. These guidelines consist of eight principles which have influenced many later lists of principles.[16] It is possible to "map" these eight principles onto the five FTC principles to emphasize the wide area covered by the FTC's principles and to show that those five principles are sufficient to cover many privacy concerns. The principles of the OECD consist of a collection limitation principle; a data quality principle; a purpose specification principle; a use limitation principle; a security safeguards principle; an openness principle; an individual participation principle; and an accountability principle.

The collection limitation principle calls for limits to the personal information collected. It also requires that the subject be informed of the collection, and when possible that consent be obtained from the subject. This principle thus falls under the FTC's notice/awareness principle. The data quality principle requires that collected data be kept accurate and up-to-date, and so it falls under the access/participation principle. The purpose specification principle requires that information gathered must have the purpose for its collection and use specified at the time of

gathering, and the information must be put to no other use. This falls under the security/integrity principle. The use limitation principle requires that consent be obtained from the data subject should the data be used for other purposes than those agreed to, or should the data be disclosed to others. This falls under the choice/consent principle and also the security/integrity principle. The openness principle requires that information collectors be open about their policies, practices, and developments with respect to personal information. It also requires that the collector respond to inquiries from data subjects and identify themselves. This falls under access/participation and notice/awareness. The security safeguards principle requires that data be protected from loss and unauthorized access, use, or disclosure. This falls under the security/integrity principle of the FTC. The individual participation principle provides the subjects of data collection with the right to know whether or not data has been collected about them; to access the data about them; and to challenge and correct data about them. This falls under the access/participation principle. The accountability principle requires that data controllers be held accountable to these requirements, through legal or other means, and falls under the enforcement/redress principle.

It is clear, then, that the OECD principles, which were widely adopted before the emergence of the internet as it is now known, are exhaustively covered by the FTC principles.

**Personal Statements of Privacy Principles**

Beyond the principles outlined by various organizations, several statements of privacy principles made by individual librarians and information professionals are worth noting.

Although these statements do not attempt to cover every aspect of privacy and confidentiality, they do focus attention on issues which are of particular importance to librarianship.

## Featherman's Ten Commandments

John Featherman of ASIS publishes a newsletter on privacy for consumers, and at the 1997 ASIS mid-year meeting presented a set of rules for individuals to help ensure that their privacy was protected. Although many of these focus on strategies for individuals, there are several rules that are applicable to institutions as well. His rules -- he calls them the "Ten Commandments of Privacy" -- are:

1. Thou shalt keep sensitive information private.
2. Thou shalt pay in cash whenever possible.
3. Thou shalt guard thy social security number and other identification numbers with thy life.
4. Thou shalt use a paper shredder in thy daily life.
5. Thou shalt use a post office box or, better yet, a mail drop.
6. Thou shalt inspect thy credit, medical, and other personal information files often.
7. Thou shalt be circumspect in thy computer affairs.
8. Thou shalt be extremely discreet when communicating.
9. Thou shalt be diligent when choosing passwords and shalt change them often.
10. Thou shalt make a lifetime commitment to protecting you privacy.[17]

Each of these commandments is clearly directed at individuals rather than institutions, but there are several points raised here that are worth considering. The third commandment should remind us that Social Security Numbers, driver license numbers, and other personally identifying numbers must be guarded with special attention. Many libraries never request this kind of information, but some may use them in patron records. For example, the Kent State University

library uses student ID cards as library cards, and the student number printed and encoded on each card is usually the individual's Social Security Number. The third commandment should call attention to the sensitivity of these student ID numbers.

The fourth commandment is important at an institutional level because paper records are often produced in libraries, as in the ILL and book request process. Computer print-outs of web pages produced inside a library may also be potentially identifiable, because the most popular browsers include the time and date as headers or footers in print-outs, and those libraries that require registration to use computer terminals make it possible for a persistent and industrious snoop to identify the patron who printed a particular page. For this reason, abandoned print-outs should be destroyed and not merely thrown out.

Commandment six calls on individuals to exercise their right to access information collected about themselves, as outlined in the FTC's principle of access. Commandment nine calls attention to the need for librarians to periodically change the passwords that provide access to personal information -- and the principle of security holds that we ought to restrict access to personal information with passwords.

Finally, commandment seven focuses on computers as potential privacy hazards. Featherman explains that encryption should be used whenever personal information is transmitted, and we should add that libraries would do well to use encryption whenever they transmit information that may identify patrons, such as ILL requests.

## Gorman's "New Libraries, Old Values"

Michael Gorman lists privacy as one of eight values of that underpin librarianship. He writes:

> The confidentiality of library records is not the most sensational weapon in the fight for privacy, but it is important, both on practical and moral grounds. In practical terms, a lot of the relationship between a library and its patrons is based on trust, and, in a free society, a library user should be secure in trusting us not to reveal what is being read and by whom. On moral grounds, we must start with the premise that everyone is entitled to freedom of access, freedom to read texts and view images, and freedom of thought and expression. None of these freedoms can survive in an atmosphere in which library use is monitored and individual reading patters are made known to anyone without permission.[18]

Gorman quite rightly focuses attention on the role of trust in library patronage, and speaks in technologically neutral language about the right to read text and view images with confidence that one's information seeking will not be revealed to others. In the context of online information seeking, this means that a user should never be identified with web pages or computer files accessed.

## Coyle's Elements of a Privacy Policy

Karen Coyle, a privacy advocate and member of the ALA's Task Force on Privacy and Confidentiality in the Electronic Environment, has outlined essential elements for a library web site privacy policy.[19] Her elements fall under four main categories: Background Information, Your Library's System, Policy Elements, and Links.

Background information should include both legal and policy statements. Coyle recommends providing information about state and local laws concerning public records and library records. This seems reasonable, as users may find it reassuring to know what protections exist. But they will also find it useful to know what limits there may be in the protection of confidentiality where laws diverge from library values. The policy statements would include both the ALA's policy on confidentiality and local policies -- that is, the policy of the institution to which the library belongs and the specific policies of the library.

The section on your library's system would outline the specific issues raised by the network, web page design, and connected but remote systems. These system elements would include a variety of matters. The logging system the library uses on its web server, which tracks http requests and browsing, should be explained. Likewise, any personalization features on the system should be explained so that users would know what information about them is stored in the system, such as their log-in history, e-mail features, and document delivery. Finally, the user should be made aware that personally identifying information may be requested by remote databases accessed through the library system or furnished to interlibrary loan partners.

The policy elements would contain details about the access, integrity, and security of data stored on the library system. The protections in place, the length of retention, and who can access the data should all be disclosed.

Lastly, the policy should include links to further clarify and explain privacy issues. Coyle specifically recommends that the online ALA policy manual should be linked, as well as the ALA task force report on online privacy. This list of links should not be considered exhaustive, but note that protecting user privacy and providing privacy literacy training can be

distinguished. It is beyond the scope of my project to create a privacy literacy curriculum.

## III. SPECIFIC PRIVACY ISSUES OF THE ONLINE ENVIRONMENT

### Cookies

A "cookie" is small text file -- 4k or smaller -- that records information on the hard drive of a computer requesting a web page. The content of a cookie generally consists of three main elements: the address of the server or web site which sent the cookie, an identification code for the specific computer, and a set of codes recording information about the user browsing habits and/or preferences. The contents of the cookie are encoded so that the recipient is unlikely to be able to interpret what they actually contain. Here is an example cookie from my own computer, the filename of which is "anyuser@fedex[1].txt":

```
FedEx_accrue
12000206511118www4744177
fedex.com/
0
2758148096
31550824
1850881248
29380962
*
```

The above contents are meaningless to me, although I can recognize the server domain "fedex.com." Presumably the codes are used by FedEx to personalize their web pages when I view them, or could be used to keep track of goods or services I requested. But in order to protect the privacy of web surfers, there ought to be both notice that this information is collected, and some way to insure that the information is accurate, if this information is in any way

identifiable with the surfer.

It should be noted that some cookies will use potentially identifying user names in the filename. For example, many cookies on my computer have filenames like "mike@nytimes.com." The cookie used my Windows logon name, which is "mike," as my userid. Because many computer networks require that users log on with some name, this method of naming cookie files presents a minor privacy threat. Many web servers will maintain databases of users and their browsing habits, and these userids are sometimes based on the user's actual name ("mmonaco" for example). Any such cookies certainly seem to constitute a personally-identifying record, even if the identification is would require a bit of sleuth work to connect it to the actual user.

There are two kinds of cookies: session and persistent. A session cookie is deleted by the browser as soon as the cookie sending server is exited. A persistent cookie will remain on the hard drive for a specified period of time. The only difference between session and persistent cookies is that persistent cookies, when set by the server, have an additional variable in the command which sends the cookie: "expires=date."[20] Session cookies are somewhat less threatening to privacy than persistent cookies because they are deleted after every session, thus limiting the amount of information a particular web site collects about the user. However, because many servers will keep their own databases of user information, there is still a need to disclose the use session cookies to users in order to protect their privacy.

Most commercial web sites and many non-commercial web sites use cookies. This is because there are many legitimate uses for cookies. Cookies which track the browsing habits of users provide information useful for improving and designing web sites. Cookies can also add

several features to web sites, such as custom formatting of web sites (some users may request specific colors, fonts, or font sizes for easier reading), alerts about new material on the site (a persistent cookie would record the last visit of a user and allow the server to provide a list on material since that last visit), and "shopping baskets" to keep a record of items requested, online forms completed, and the like.[21] Disabling cookies (an option provided by browsers which use them) can therefore create a loss of functionality, and some web sites would be unusable without cookies.[22]

Cookies present several privacy issues to users of online resources. At the most basic level, cookies may be considered a privacy threat because they can operate as "spies," in the sense that the senders of cookies may not reveal themselves while at the same time collecting information about the user. An informal study by Karen Coyle of cookies collected on her hard drive revealed that about one third of the cookies were untraceable, owing to the fact that the web addresses in the cookies could not be reached.[23] The cookie senders thus remain anonymous, which in itself raises questions about the intentions of the cookie senders.

It has been noted that cookies, by themselves, do not actually identify the user who receives them. Cookies simply collect clickstream data, indicating what a user views, how long he or she is connected to the sites, and which advertisements were displayed by the site. By itself, this data does constitute useful information for the web site designer and marketer. The real privacy threat emerges when cookie technology is combined with another technology: online forms. Users may be asked for identifying personal information including addresses, names, phone numbers, and the like when they register to use various web sites, services, or files. If this information should be matched with clickstream data collected from cookies, the user has

suddenly revealed much more than the information voluntarily released in the form. A profile of the users' interests can be determined by studying the clickstream data from cookies which reveal where, when, and for how long the user has surfed web sites. The appeal of such information for marketers and advertisers is obvious. But the users' privacy has been compromised if the information collected by cookies is collected covertly, without disclosure.

It is believed that most cookies can be accessed only by the site which created them. However, a bug in Microsoft's Internet Explorer was discovered in May of 2000, called the "Open Cookie Jar." This bug allowed any server to read any cookie, provided that the name of the site which created the cookie was known.[24] Microsoft has, since then, corrected the problem, and even offers cookie management features in the newest version of Internet Explorer. Even so, this episode calls attention to the potential danger that cookies be intercepted or viewed illicitly. A similar bug was found in one version of Netscape Navigator. Bugs such as these clearly violate the principle of security, as they allow unauthorized disclosure and use of the information in the cookie.

It is possible for the server which sends cookies to specify that cookies will be sent only over secure communication channels; the set-cookie http response header would need to include the command "secure." Presently, only HTTPS servers (http servers on SSL) are considered "secure" by the Netscape browser.[25]

One further privacy threat posed by cookies lies in the fact that some banner advertisers send cookies as well. Because the server providing banner ads sets the cookie rather than the server providing the actual page visited, these banner ad companies are able to track a user's browsing across many different sites and servers.[26] The purpose of these cookies is to aid in

targeting specific ads at users. However, because these companies store information about the recipients of the banner ads and their browsing habits, the user's privacy is seriously threatened, especially if this information is combined with forms which ask for personal information. It is one thing if Doubleclick.com tags my computer with a cookie that records my browsing habits as "mike@doubleclick.com" and quite another if they should be able to connect my userid with my name and address because I filled in a form to register to use the online New York Times site.

Libraries should consider cookies carefully. Computers provided for patron use will accrue a large collection of cookies. The risk to patron privacy lies in the danger that when a patron fills in an online form, their personal information may be erroneously linked to the browsing habits of other users. But personally identifiable information must be accurate, according to the principle of integrity discussed above. On the other hand, patrons may be accurately profiled and their privacy compromised if they are unaware that the library's computers accept and use cookies. There is also the issue of library web sites using cookies. Better access may be provided to users by allowing them to specify larger fonts or other formatting issues. And librarians would be able to improve their web sites by reviewing usage. But these benefits must be balanced against the danger that cookies from library web sites could be intercepted or accessed illicitly. The danger of interception may be reduced by always specifying that cookies be sent only over secure communication lines. The danger that others may view cookies, for example by exploiting the Open Cookie Jar bug, is probably best met by using only session cookies.

**Clickstream and Log Files**

Apart from cookies, there are other traces of internet use recorded by web servers and private firms. These other records are loosely grouped under the terms "clickstream," "mouse droppings," and "data trail." This discussion will use the term "clickstream" as it is the most widespread term.

One source of clickstream data is found in the headers sent by web browsers to servers when a user accesses a web page. These headers reveal the internet protocol address ("IP address") of the user, the time of day when the page is accessed, the pages and images downloaded, the page from which the user is accessing the new page ("Referrer"), and data entered by the user into forms. Cookies are sometimes considered to be part of this clickstream data as well, but cookies will not be explored further here.

In order to give a picture of what one might look like, a sample log file from Karen Coyle's web pages is reproduced here:

ppp156210.asahi-net.or.jp -- [29/Mar/1998:05:45:00 -0800] "GET /~kec/best.html HTTP/1.0" 200
DIAL8-ASYNC38.DIAL.NET.NYU.EDU -- [29/Mar/1998:12:23:11 -0800] "GET /~kec HTTP/1.0" 302
gwc-mel.afwgc.af.mil -- [29/Mar/1998:14:24:09 -0800] "GET /~kec/ HTTP/1.0" 304[27]

As can be seen, an IP address is recorded, and then the date and time that pages are accessed, and finally the path of the page accessed. Each of three entries follows the format "IP address -- date/time -- command/request."

All of this data is provided by the web browser used to access the pages. But the

individual using the browser is likely to be unaware that this information was provided to the server visited. The consumer advocacy web site "Junkbusters.com" has a page which allows users to see for themselves what their browser tells every server they visit. When visiting the Junkbusters page with a computer in a KSU philosophy department, the following output is produced:

> The ``HTTP Referrer" tells them what led you to the request. In this case it was **http://www.junkbusters.com/junkdata.html**.
> Some advertisers choose which ad to show you based on two variables that tell them which computers you are coming through: the ``Remote Address" (in your case 131.123.68.123) and the ``Remote host:" (in your case **131.123.68.123**). These variables may give indications of where you live or work.
> The ``User Agent" variable, in your case **Mozilla/4.73 [en] (Win95; U)**, indicates what software and hardware you are using. This interests companies that sell these goods. (Mozilla is Netscape's name for its browser.)
> Some browsers actually hand over your email address or other indications of your identity. In your case it appears the ``HTTP From" variable was **not provided**, and the ``REMOTE_USER" variable was **not provided**.[28]

This URL basically displays the headers sent by the browser when a user visits the address -- they are displayed above in bold face. The text on the page also explains the specific variables used by http to record this information. Junkbusters encourages people interested in privacy to link to this page in order to promote privacy literacy (and to promote their privacy software). Libraries would do well to make this page available to their users. But perhaps the most interesting thing to notice here is that there are header variables not recorded in Karen Coyle's sample log entries but which other servers might choose to log.

Another source of clickstream data is in the browser "add-ons" known as Java, JavaScript, and VB Script. These are all technologies which enhance the web pages with graphics, animation, and the like. These will generally provide information about the browser

being used, screen resolution, and other data of concern mainly for formatting purposes.[29]

The information sent back and forth in headers is of little concern on its own, but a threat to privacy emerges when this information is stored for later use and analysis in transaction log files.

One thing to notice is that the clickstream data stored in log files would identify computers but not individuals.[30] However, an IP address can reveal a certain amount of information about an individual. IP addresses are assigned by the internet service provider (ISP), and often disclose the region of the country or even the city the computer is in. A quick review of the IP addresses in Karen Coyle's log file suggests that the three transactions logged were requested by users in Japan, New York University, and a military base. IP addresses that consist entirely of numbers are in principle traceable as well. The IPs also reveal the identity of the ISP, which in itself reveals little about the individual. The real danger emerges when clickstream data is merged with personally identifying information. If your identity can be connected with your clickstream, others can know what web sites you visit, which links you follow, what files you download or view, and so on.

Perhaps the most egregious and high profile case of compromising privacy in this manner is the recent controversy surrounding DoubleClick.com. DoubleClick.com acquired Abacus Direct, a company which maintains databases of consumer information compiled by direct marketers. The kind of information Abacus Direct maintains involves consumer habits and demographics. DoubleClick.com provides banner ads to thousands of web sites, and as described above, these banner ads send cookies to web surfers who see their ads. These cookies identify the user to DoubleClick.com, and the browser headers transmitted when surfers click on

the banners are compiled in DoubleClick's log files. When the Abacus Direct databases are combined with DoubleClick.com's log files, the resulting database identifies individuals with demographic information, names and addresses, web surfing habits, and consumer information. All of this information has obvious use for advertisers who want to target ads effectively. But this information also constitutes a severe compromise of the surfer's privacy, not least because the individual is unlikely to be aware of the clickstream information collected.

A persistent problem with online privacy is the fact that so many users are unaware of the information they are providing to web sites and servers. Online forms represent the only conscious, voluntary release of information for most web users. The average web user simply can't know what cookies to accept or reject, and almost certainly does not realize a trail is created whenever he or she clicks on a link.[31] Again, the principle of notice requires that the surfer be notified that header information is logged.

DoubleClick.com, in the face of criticism from consumer advocates and privacy advocates, has promised to await the development of internet privacy rules before merging the log files and databases. But it is evidently just a matter of time before they or other companies will begin to create detailed profiles of web users.

## Email

Because some library patrons use email to ask reference questions, and because some libraries offer internet access which is used by patrons to access email accounts, email policies

are often adopted by libraries or their parent institutions. These policies should include a privacy element for several reasons.

First, the Electronic Communications Privacy Act, passed by Congress in 1986, guarantees certain protections from the interception and disclosure of email messages. Just as the confidentiality polices of the ALA call on libraries to protect patron records and disclose them to law enforcement or other government agencies only in cases were good reason and proper legal documentation are provided, the email messages of library patrons must be protected.[32]

But many institutions place limits on the appropriate use of email accounts, and libraries may do the same. The crucial point is that any policies must be disclosed to users. So if the library monitors email messages created on library computers or sent to reference personnel, these practices must be disclosed in the privacy policy statement.

Lastly, because emailed reference questions are created by patrons in the information seeking process, they constitute records that ALA policies require libraries to protect as confidential.

In order to adequately protect user confidentiality, then, the online privacy polices of libraries must disclose any monitoring of email. Presumably, any such monitoring would be undertaken only in order to assure that library resources are not abused or as part of quality assurance in the case of reference questions. The library should not keep unnecessary records of messages sent and received, and protect those it records as confidential.

**Interlibrary Loans**

Requesting and providing materials for interlibrary loans (ILL) is another library activity which generates records. These records are likely to identify the patron requesting materials. Christopher Nolan, writing in the Journal of Academic Librarianship, notes that the paper ALA-sanctioned ILL request form and the online work forms used by OCLC, RLN, and WLN all use information about the requesting patron, and supply this information to the institution making the loan.[33] This creates a privacy issue for patrons because both the institution making the request on their behalf, and the lending institution, have records of the patron's information seeking.

There is obviously good reason to keep track of circulation, but Nolan points out that several things may happen to the ILL request. Because the patron is identified in ILL requests, traces of the loan will remain even if the library conscientiously destroys old circulation records. First, the CONTU guidelines attached to copyright laws require that records of all photocopies of materials obtained through ILL be maintained for three years. Commonly, these records are maintained by keeping a copy of the ILL form which generated the photocopy.[34] Because the patron is identified on this form, there is now a record of materials borrowed which is identifiable with the patron.

Second, the ILL request forms are supplied to the lending institution. Again because the forms contain the patron's name and possibly other information, the lending institution now has a record of the patron's information seeking.

Privacy policies must address these records. An ideal policy would require that CONTU guidelines be met by keeping a non-identifying record of photocopies made and would require

that patron information not be sent to the lending institution. Because some ILL forms are online, it would be a good idea to disclose the library's policies about ILL requests in the online privacy policy.

## Searches

Another privacy issue of the online environment is internet searching. There are two ways that searching the internet can threaten a library patron's privacy. First, many search engines and directory sites will record the actual search terms entered by the user in the URL of search results. For example, doing a simple search of the term "clickstream" at Yahoo and Excite lead my browser to the following addresses:

http://search.yahoo.com/bin/search?p=clickstream
http://search.excite.com/search.gw?search=clickstream

If I were to follow the links provided on this page, my browser's header would disclose this URL as the "HTTP Referrer" variable. Thus, my search strategy is potentially recorded in a log file, and it is connected with my IP address. If a library web site provides links to search engines which encode search terms in the results URL, this should be mentioned in the privacy policy.

Second, many search engines use cookies to keep track of users and their queries. This is not a privacy problem when the search engine keeps only "aggregate" information -- that is, statistics about user trends and surfing habits. But persistent cookies used by search engines to track a particular user's queries and searches do threaten privacy. As discussed above under the

topic of cookies, there is always a risk that cookie information will be linked with the subject's identity when the subject completes a form or otherwise releases information online. And as in most online privacy issues, users may be unaware that information is being disclosed.

Apart from internet searches, libraries need to consider the possibility that OPACs can retain digital traces of the searches conducted by users. Users occasionally leave a terminal with their search results and strategy visible on the monitor. In so far as the patron has chosen to leave the terminal without returning to the menu screen or otherwise concealing his or her search, there is little the library can do to protect the confidentiality of searches from this kind of disclosure. But libraries will also have to consider the possibility that the OPAC retains records of previous searches in its memory, disk cache, or log files. Whenever this is the case, the policy of the library must be to ensure that these records are secure from unauthorized access and destroyed as soon as they serve no role in the library's functioning. At the very least, records of searches maintained for statistical analysis should have no personally identifying information.

In most cases, OPACs do not require logging in or identifying the user. But there are exceptions, most notably in systems which allow users to view their own circulation records. Because this confidential information is accessible, it is the library's responsibility and duty to take adequate measures for protecting the information. Circulation records should be protected by passwords. They should also be transmitted to authorized accessers in an encrypted format to protect against interception.

## IV. CONCLUSIONS

The protection of privacy can be assured by meeting the five main principles outline by the FTC: notice/awareness, choice/consent, access/participation, integrity/security, and redress/enforcement. Librarians must strive to uphold each of these principles in their handling of confidential information.

The information generated by online library use includes cookies, clickstream data (including the header data created by browsers and log files maintained by servers), ILL request forms, and email reference questions. All of this information is potentially personally identifiable and thus should be protected by strict standards of privacy as outlined in the five principles. When data is abstracted from log files into statistical data or otherwise severed from personally identifying marks, the data can be considered outside the umbrella of privacy protection.

Insofar as library web sites and online systems collect personally identifying information or provide it to outside agencies such as online search engines or databases, there must be a clear disclosure of this practice and adequate measures taken to protect patron confidentiality.

Because libraries also provide access to databases, search engines, and web pages which belong to other agencies, it is also important that libraries give notice to the patron that the privacy policies of the library are not shared by these outside agencies. The patrons should be reminded that they cannot assume resources accessed through the library are covered by the

same guarantees of privacy that library makes about use of resources belonging to it.

All disclosures of policies should be made in a privacy policy statement that can be accessed easily by online patrons, and an ideal policy will consider all of the issues outlined in this paper.

## NOTES

[1] Available at URL: http://www.ala.org/alaorg/policymanual/libperson.html

[2] Available at URL: http://www.ala.org/alaorg/policymanual/libserve.html

[3] American Library Association Office for Intellectual Freedom, pp. 150-158.

[4] Available at URL: http://staffweb.library.vanderbilt.edu/ala_tf/report.htm

[5] American Library Association Task Force on Privacy and Confidentiality in the Electronic Environment. Available at URL: http://staffweb.library.vanderbilt.edu/ala_tf/report.htm

[6] Federal Trade Commission,1998, p. 7.

[7] Federal Trade Commission, 2000, p. iii.

[8] FTC, 1998, p. 7.

[9] FTC, 1998, p. 7-8.

[10] FTC, 1998, p. 8.

[11] Ibid.

[12] FTC, 1998, p. 9.

[13] FTC, 1998, p. 10.

[14] Ibid.

[15] Ibid.

[16] Center for Democracy and Technology, "The OECD guidelines." Available at URL: http://www.cdt.org/privacy/guide/oecdguidelines.html

[17] John Featherman, "Moses meets Big Brother: the ten commandments of privacy." Available at URL: http://www.asis.org/bulletin/Feb-97/featherman.html

[18] Michael Gorman, p. 49.

[19] Karen Coyle, "Elements of a library web site privacy policy." Unpublished document.

[20] The syntax of cookie specification for the Netscape browser is described in some detail at URL: http://home.netscape.com/newsref/std/cookie_spec.html . The specifications for Microsoft's Internet Explorer are similar.

[21] Interlog Internet Services. "Cookies: what are they, and why do you want to give me one?" Available at URL: http://www.interlog.com/cookies.html

[22] Ibid.

[23] Karen Coyle, "A Cookie Study." Available at URL: http://www.kcoyle.net/cookie_study.html

[24] Kevin Townsend, "Briefing Paper: Why cookies cause upsets." Available at URL: http://www.zdnet.co.uk/itweek/brief/2000/32/internet/

[25] Netscape Support Documentation. "Persistent Client State HTTP Cookies." Available at URL: http://home.netscape.com/newref/std/cookie_spec.html

[26] Interlog Internet Services.

[27] Karen Coyle, "A Primer on Internet Privacy." Available at URL: http://www.kcolye.net/privacyprimer.html

[28] This page is available at URL: http://junkbusters.com/cgi-bin/privacy, and provides personalized outputs for each computer which accesses it.

[29] Center for Democracy and Technology. "Online tracking FAQ." Available at URL: http://www.cdt.org/privacy/guide/start/track.html

[30] Coyle, "A Primer."

[31] Ibid.

[32] Pat Gannon-Leary, pp. 221 - 225.

[33] Christopher Nolan, p. 82.

[34] Nolan, p. 83-84.

## WORKS CITED

American Library Association. ALA Policy Manual. Online Document. Available at URL: http://www.ala.org/alaorg/policymanual/index.html; accessed 11/8/00

American Library Association Office for Intellectual Freedom. Intellectual Freedom Manual. Chicago: The Association, 1996.

American Library Association Task Force on Privacy and Confidentiality in the Electronic Environment. Report of the Task Force on Privacy and Confidentiality in the Electronic Environment: Final Report, July 7, 2000. Online Document. Available at URL: http://staffweb.library.vanderbilt.edu/ala_tf/report.htm; accessed 11/8/00.

Center for Democracy and Technology. "The OECD guidelines." Online Document. Available at URL: http://www.cdt.org/privacy/guide/oecdguidelines.html; accessed 11/12/00.

Center for Democracy and Technology. "Online tracking FAQ." Online Document. Available at URL: http://www.cdt.org/privacy/guide/start/track.html; accessed 11/12/00.

Coyle, Karen. "A Cookie Study." Online Document. Available at URL: http://www.kcoyle.net/cookie_study.html; accessed 11/3/00.

Coyle, Karen. "Elements of a library web site privacy policy." Unpublished document.

Coyle, Karen. "A Primer on Internet Privacy." Online Document. Available at URL: http://www.kcolye.net/privacyprimer.html; accessed 11/3/00.

Featherman, John. "Moses meets Big Brother: the ten commandments of privacy." Online Document. Available at URL: http://www.asis.org/bulletin/Feb-97/featherman.html; accessed 1/18/01.

Federal Trade Commission. Privacy Online: A Report to Congress, 1998. Online Document. Available at URL: http://www.ftc.gov/reports/privacy3/priv-23a.pdf; accessed 1/11/01.

Federal Trade Commission. Fair Information Practices in the Electronic Marketplace, 2000. Online Document. Available at URL:

http://www.ftc.gov/reports/privacy2000/privacy2000.pdf; accessed 1/11/01.

Gannon-Leary, Pat. "'E' for Exposed? E-mail and Privacy Issues." The Electronic Library, 15(3), pp. 221 - 225.

Gorman, Michael. "New Libraries, Old Values." The Australian Library Journal, Feb. 1999, pp. 43-52.

Interlog Internet Services. "Cookies: what are they, and why do you want to give me one?" Online Document. Available at URL: http://www.interlog.com/cookies.html; accessed 1/9/01.

Junkbusters. "Junkbusters Alert on Internet Privacy." Online Document. Available at URL: http://junkbusters.com/cgi-bin/privacy; accessed 11/12/00.

Netscape Support Documentation. "Persistent Client State HTTP Cookies." Online Document. Available at URL: http://home.netscape.com/newsref/std/cookie_spec.html; accessed 1/18/01.

Nolan, Christopher. "The confidentiality of interlibrary loans." The Journal of Academic Librarianship, 19(2), pp. 81-86.

Townsend, Kevin. "Briefing Paper: Why cookies cause upsets." Online Document. Available at URL: http://www.zdnet.co.uk/itweek/brief/2000/32/internet/ ; accessed 1/26/01.

# APPENDIX ONE: THE MODEL POLICY

## A Model Privacy Policy for Library Web Sites

Because libraries may vary considerably in their policies and practices, some provisions of this policy are bracketed, indicating that the content will vary according to local practices. Bracketed text marked "comment" provide suggestions for personalizing the policy rather than suggested provisions. Sentences marked with an asterisk (*) are appropriate for many libraries but may also vary according to local practices.

### Privacy Policy

**Contents**

General principles
The scope and limits of this policy
What information is collected
Your choices about information collection
How to access the information we collect about you
The accuracy and security of this information
What you can do if you want to know more
Links to privacy information and library policy pages

**General principles**

1. The policy of the American Library Association (ALA)

As information professionals, we uphold the ALA's ethical code regarding privacy . We strive to protect each user's privacy and confidentiality, and toward this end, we adopt the this policy.

"We extend confidentiality to all information sought or received, materials consulted, borrowed, acquired, or transmitted, including database search records, reference interviews, interlibrary loan records, and other personally identifiable uses of library resources, services, and materials. Any record identifying a library user with specific materials will be regarded as confidential and protected. Such records will not be made available to any outside agency, institution, or private firm, or any government agency except pursuant to such process, order, or
subpoena as may be authorized under the authority of, and pursuant to, federal, state, or local law relating to civil, criminal, or administrative discovery procedures or legislative

investigatory power. We will resist the issuance or enforcement of any such process, order, or subpoena until such time as as a proper showing of good cause has been made in a court of competent jurisdiction." (ALA Policy Manual)

2. The policy of affiliated institutions and/or schools, and the policy of this library [each library, as part of some larger institution or school, will need to consult the policies of the institution or school. Also, if the library has any specific privacy policies of its own, apart from those of umbrella organizations and the like, will need to articulate them here.]

**The scope and limits of this policy**

1. This policy is intended to extend to physical records generated in the information seeking behaviors of all library patrons. Such physical records include papers such as written Interlibrary Loan requests (whether written by the patron or library staff), written requests for items held on reserve or in special collections, and any notifications of overdue items or concerning the availability of requested items.

2. This policy is also intended to extend to electronic or digital records generated by the use of library resources and resources provided by but not belonging to the library. Such digital records include search strategies on library OPACs, databases, and search engines; digital
requests for materials; electronic requests for information such as emails requesting reference help; and data collected by library web sites either explicitly (as when forms request information) or implicitly (as when our server requests information from your web browser).

Specific polices regarding electronic library resources and online library use. The right of privacy includes five subsidiary rights: notice/awareness , choice/consent , access/participation , security/integrity , and redress/enforcement . That is, you have a right to know what information is collected, a right to choose if you want information to be collected, a right to review any such information to ensure its accuracy, and a right to have any such information protected from unwarranted or illegitimate use. You also have a right to enforce the preceding rights and seek redress through the appropriate channels.

**Notice/Awareness**

The **library** collects information about patrons in several ways.

- In order to obtain a library card for checking out books and other materials, patrons must supply the library with their name, address, and phone number.*
- Interlibrary loan requests -- requests for materials not owned by the library -- may require the same information which is furnished to the loaning library.
- [Patrons using the library's computer terminals are asked to supply their names and the time of day of usage.]

All of this information is considered confidential by the library and will not be disclosed to any outside agencies or persons, except in case where a justified and bona fide subpoena is submitted. The library maintains permanent records of patrons but destroys all records of information seeking (including circulation records and ILL requests) upon return of the materials.

The **library's web site** also collects information about patrons in a variety of ways.

All visitors ot this web site have information collected about them automatically. This information includes:

- the internet domain from which this site is accessed
- the internet address ( IP address ) from which this site is accessed
- the type of browser and operating system you are using on your computer
- the date and time you visit
- the pages you view while here
- if you followed a link from some other site to get to this site, the address of that web site
- any searches conducted on this site, including both the search terms you use and the links you follow

Most of this information is supplied automatically by your web browser. This is called header information. You can see exactly what information your browser supplies to this and every other web site you visit by following this link: www.junkbusters.com/cgi-bin/privacy. This information is used by the library's web site maintainer to make the site easier to use and more efficient. The library uses this information to learn about the habits of visitors and what areas of the web site they use the most. This information is never stored in a record which can identify particular users and what they access -- instead, only aggregate, statistical records are kept.

- Information about your browsing on this site, such as which pages you view and how long you view them, are stored in digital files called log files. These log files will never identify particular users and is not shared with any outside agencies.
- This web site does not use cookies .* [If cookies are used , this provision should identify the kind of information stored in them, the type of cookies used -- persistent or session , and why they are used.]

This web site has links to databases, search engines, and other web sites which may collect personal information about visitors. We advise all patrons to refrain from disclosing any personal information on the internet unless the patron is reasonably certain that this information will kept secure and confidential. Patrons should also be aware that these sites may keep records of any searches performed, web sites visited, and files

downloaded. Patrons should be aware that such information may be stored in records that can identify the patron to these outside agencies, and that these agencies may in turn share or sell this private information. For more information about privacy online, follow these links:
ALA Task Force Report on Privacy in the Online Environment
A Primer on Internet Privacy
The Electronic Privacy Information Center
Surfer Beware: Personal Privacy and the Internet
[comment: these links are provided to patrons as useful starting points for privacy literacy.]
For more information on the American Library Association's policies, see The ALA's Policy Manual

**Choice/Consent**

The library must collect certain information in order to provide services. Patrons may choose to opt-out of this information collection, but this may affect the services the library provides. Specifically, the library materials can be checked out of the library only by patrons who have released their names, addresses, and phone numbers,* as the library needs some way to contact patrons should the materials become overdue. Any other information collection, not connected to services provided by the library, will always be on an "opt-in" basis -- that is, the patron initiates the collection of information on a voluntary basis.

Information collected by the library web site is, in general, never connected to the user. The information collected is automatically furnished by the user's web browser. The library will always ask the patron for permission to collect any information that can be identified with the patron. [ for libraries with web sites that use cookies: The patron can control whether or not cookies will be stored on his or her computer, by adjusting the settings of their browser. For Microsoft Internet Explorer users, this option can be found under the "Tools : Internet Preferences" pull-down menu; for Netscape, this option is found under the "Edit : Preferences" menu.]

**Access/Participation**

All library patrons who have had personally identifying information about them collected have right to be sure the information is accurate and up to date. All information collected in connection with the issue of library cards may be reviewed and updated by contacting [comment: insert the appropriate administrator and contact details; this will presumably be the same administrator mentioned below ]. Patrons may view their current circulation records and should see a librarian if they find any inaccuracies. [comment: If circulation records are available for online viewing, instructions for accessing them should be placed here.]

**Security/Integrity**

The library takes the integrity and security of all records seriously. All records which can be identified with a patron are stored on computer systems which are not accessible by agencies or individuals outside the library, and only authorized personnel have access

to such records inside the library. All of the library's records which are kept on computers which can be accessed from outside via the internet are protected by security software.* [comment: the details here will vary and should be disclosed in a manner than that does not itself pose a security risk]

The library will update patron records as needed to ensure that all data is accurate. The library does not share any patron information with other government or private agencies.

**Redress/Enforcement**

The library recognizes that privacy polices must be enforced and will take appropriate measures to redress any violation of a patron's privacy. Please contact [the administrator in charge of policy and security] if you have any questions about what personally identifing records about you are maintained or if you feel that your confidentiality has been compromised.

[comment: There should be every effort to make this administrator accessible to the public for questions and assurances, through email, phone, and personal contact. This administrator should oversee the enforcement of library policies and keep the staff informed about these policies.]

[comment: There should also a brief disclosure of any state or local laws which oversee the confidentiality of library records or potentially overrule these polices, such as "sunshine laws."]

**Links**

The ALA's Policy Manual
ALA Task Force Report on Privacy in the Online Environment
A Primer on Internet Privacy
The Electronic Privacy Information Center
Surfer Beware: Personal Privacy and the Internet
Junkbusters' Privacy Alert Page

## APPENDIX 2: HTML SOURCE FOR THE MODEL POLICY

```
<!doctype html public "-//w3c//dtd html 4.0 transitional//en">
<html>
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
  <meta name="Author" content="Michael Monaco">
  <meta name="GENERATOR" content="Mozilla/4.73 [en] (Win95; U) [Netscape]">
  <title>MODELPOLICY</title>
</head>
<body>
<br>
<center>
<p><b><font size=+1>A Model Privacy Policy for Library Web Sites</font></b></center>

<p>     Because libraries may vary considerably in their policies and
practices, some provisions of this policy are bracketed, indicating that the content will vary
according to local practices.  Bracketed text marked "comment" provide suggestions for
personalizing the policy rather than suggested provisions.  Sentences marked with an
asterisk (*) are appropriate for many libraries but may also vary according to local practices.
<p><b><font size=+1>Privacy Policy</font></b>
<p><b>Contents</b>
<br>
<br> <a href="#Principles">General principles</a>
<br> <a href="#Scope">The scope and limits of this policy</a>
<br> <a href="#Notice">What information is collected</a>
<br> <a href="#Choice">Your choices about information collection</a>
<br> <a href="#Access">How to access the information we collect about you</a>
<br> <a href="#Security">The accuracy and security of this information</a>
<br> <a href="#Redress">What you can do if you want to know more</a>
<br> <a href="#Links">Links to privacy information and library policy pages</a>
<br> 
<p><a NAME="Principles"></a><b>General principles</b>
<p>    1.  The policy of the American Library Association
(ALA)
<br>    As information professionals, we uphold the <a
href="OnlinePrivacy.htm#alapolicies">ALA's ethical code regarding privacy</a> .  We
```

strive to protect each user's privacy and confidentiality, and toward this end, we adopt the this policy.

<p>    "We extend confidentiality to all information sought or received, materials consulted, borrowed, acquired, or transmitted, including database search records, reference interviews, interlibrary loan records, and other personally identifiable uses of library resources, services, and materials.  Any record identifying a library user with specific materials will be regarded as confidential and protected.  Such records will not be made available to any outside agency, institution, or private firm, or any government agency except pursuant to such process, order, or

<br>subpoena as may be authorized under the authority of, and pursuant to, federal, state, or local law relating to civil, criminal, or administrative discovery procedures or legislative investigatory power.  We will resist the issuance or enforcement of any such process, order, or subpoena until such time as as a proper showing of good cause has been made in a court of competent jurisdiction."  (ALA Policy Manual)

<p>    2. The policy of affiliated institutions and/or schools,

and the policy of this library

<br>    [each library, as part of some larger institution or school, will need to consult the policies of the institution or school.  Also, if the library has any specific privacy policies of its own, apart from those of umbrella organizations and the like, will need to articulate them here.]

<p><a NAME="Scope"></a>The scope and limits of this policy

<p>    1.  This policy is intended to extend to physical records generated in the information seeking behaviors of all library patrons.  Such physical records include papers such as written Interlibrary Loan requests (whether written by the patron or library staff), written requests for items held on reserve or in special collections, and any notifications of overdue items or concerning the availability of  requested items.

<p>    2.  This policy is also intended to extend to electronic or digital records generated by the use of library resources and resources provided by but not belonging to the library.  Such digital records include search strategies on library OPACs, databases, and search engines; digital

<br>requests for materials; electronic requests for information such as emails requesting reference help; and data collected by library web sites either explicitly (as when forms request information) or implicitly (as when our server requests information from your web browser).

<p>    Specific polices regarding electronic library resources

and online library use.  The right of privacy includes five subsidiary rights:  <a href="OnlinePrivacy.htm#NOTICE">notice/awareness</a> , <a href="OnlinePrivacy.htm#CHOICE">choice/consent</a> , <a href="OnlinePrivacy.htm#ACCESS">access/participation</a> ,

<a href="OnlinePrivacy.htm#SECURITY">security/integrity</a>

, and <a href="OnlinePrivacy.htm#REDRESS">redress/enforcement</a> . 

That is, you have a right to know what information is collected, a right to choose if you want information to be collected, a right to review any such information to ensure its accuracy, and a right to have any such information protected from unwarranted or illegitimate use.  You also have a right to enforce the preceding rights and seek redress through the appropriate channels.

<p><a NAME="Notice"></a>Notice/Awareness

<p>     The library collects information about patrons in several ways.
<br>
<ul>
<li>
In order to obtain a library card for checking out books and other materials, patrons must supply the library with their name, address, and phone number.*</li>

<li>
Interlibrary loan requests -- requests for materials not owned by the library -- may require the same information which is furnished to the loaning library.</li>

<li>
[Patrons using the library's computer terminals are asked to supply their names and the time of day of usage.]</li>
</ul>
     All of this information is considered confidential by the library and will not be disclosed to any outside agencies or persons, except in case where a justified and bona fide subpoena is submitted.  The library maintains permanent records of patrons but destroys all records of information seeking (including circulation records and ILL requests) upon return of the materials.
<p>     The library's web site also collects information about patrons in a variety of ways.
<p>     All visitors to this web site have information collected about them automatically.  This information includes:
<ul>
<li>
the internet domain from which this site is accessed</li>

<li>
the internet address ( <a href="OnlinePrivacy.htm#IPADDRESS">IP address</a> ) from which this site is accessed</li>

<li>
the type of browser and operating system you are using on your computer</li>

<li>
the date and time you visit</li>

<li>
the pages you view while here</li>

<li>
if you followed a link from some other site to get to this site, the address of that web site</li>

<li>

any searches conducted on this site, including both the search terms you use and the links you follow</li>
</ul>
     Most of this information is supplied automatically by your web browser.  This is called  <a href="OnlinePrivacy.htm#HEADERS">header information</a>.  You can see exactly what information your browser supplies to this and every other web site you visit by following this link:
<a href="www.junkbusters.com/cgi-bin/privacy">www.junkbusters.com/cgi-bin/privacy.</a>  
This information is used by the library's web site maintainer to make the site easier to use and more efficient.  The library uses this information to learn about the habits of visitors and what areas of the web site they use the most.  This information is never stored in a record which can identify particular users and what they access -- instead, only <a href="OnlinePrivacy.htm#aggregatenotpersonal">aggregate, statistical records</a> are kept.
<br> 
<ul>
<li>
Information about your browsing on this site, such as which pages you view and how long you view them, are stored in digital files called  <a href="OnlinePrivacy.htm#LOGFILES">log files</a>.  These log files will never identify particular users and is not shared with any outside agencies.</li>

<li>
This web site does not use  <a href="OnlinePrivacy.htm#COOKIEdf">cookies</a>
.* [If cookies are  <a href="OnlinePrivacy.htm#COOKIESuse">used</a>
, this provision should identify the kind of information stored in them, the type of cookies used --   <a href="OnlinePrivacy.htm#COOKIEtypes">persistent or session</a> , and why they are used.]</li>
</ul>
     This web site has links to databases, search engines, and other web sites which may collect personal information about visitors.  We advise all patrons to refrain from disclosing any personal information on the internet unless the patron is reasonably certain that this information will kept secure and confidential.  Patrons should also be aware that these sites may keep records of any searches performed, web sites visited, and files downloaded.  Patrons should be aware that such information may be stored in records that can identify the patron to these outside agencies, and that these agencies may in turn share or sell this private information.  For more information about privacy online, follow these links:
<br> <a href="http://staffweb.library.vanderbilt.edu/ala_tf/report.htm">ALA Task Force Report on Privacy in the Online Environment</a>
<br> <a href="http://www.kcoyle.net/privacyprimer.html">A Primer on Internet Privacy</a>
<br> <a href="http://www.epic.org">The Electronic Privacy Information Center</a>
<br> <a href="http://www.epic.org/reports/surfer-beware.html">Surfer Beware: Personal Privacy and the Internet</a>
<br>[comment: these links are provided to patrons as useful starting points for privacy literacy.]

<br>For more information on the American Library Association's policies, see  <a href="http://www.ala.org/alaorg/policymanual/index.htm">The ALA's Policy Manual</a>

<p><a NAME="Choice"></a>Choice/Consent

<p>     The library must collect certain information in order to provide services.  Patrons may choose to  <a href="OnlinePrivacy.htm#OPTIN/OUT">opt-out</a>  of this information collection, but this may affect the services the library provides.  Specifically, the library materials can be checked out of the library only by patrons who have released their names, addresses, and phone numbers,* as the library needs some way to contact patrons should the materials become overdue.  Any other information collection, not connected to services provided by the library, will always be on an "opt-in" basis -- that is, the patron initiates the collection of information on a voluntary basis.

<p>     Information collected by the library web site is, in general, never connected to the user.  The information collected is  <a href="OnlinePrivacy.htm#HEADERS">automatically furnished</a>  by the user's web browser.  The library will always ask the patron for permission to collect any information that can be identified with the patron. [for libraries with web sites that use cookies: The patron can control whether or not cookies will be stored on his or her computer, by adjusting the settings of their browser.  For Microsoft Internet Explorer users, this option can be found under the "Tools : Internet Preferences" pull-down menu; for Netscape, this option is found under the "Edit : Preferences" menu.]

<p><a NAME="Access"></a>Access/Participation

<p>     All library patrons who have had personally identifying information about them collected have right to be sure the information is accurate and up to date.  All information collected in connection with the issue of library cards may be reviewed and updated by contacting [comment: insert the appropriate administrator and contact details; this will presumably be the same administrator mentioned <a href="#admin">below</a> ].  Patrons may view their current circulation records and should see a librarian if they find any inaccuracies. [comment: If circulation records are available for online viewing, instructions for accessing them should be placed here.]

<p><a NAME="Security"></a>Security/Integrity

<p>     The library takes the integrity and security of all records seriously.  All records which can be identified with a patron are stored on computer systems which are not accessible by agencies or individuals outside the library, and only authorized personnel have access to such records inside the library.  All of the library's records which are kept on computers which can be accessed from outside via the internet are protected by security software.* [comment: the details here will vary and should be disclosed in a manner than that does not itself pose a security risk]

<p>    The library will update patron records as needed to ensure that all data is accurate.  The library does not share any patron information with other government or private agencies.

<p><a NAME="Redress"></a>Redress/Enforcement

<p>     The library recognizes that privacy polices must be enforced and will take appropriate measures to redress any violation of a patron's privacy.  Please contact <a NAME="admin"></a>[the administrator in charge of policy and security] if you have any questions about what personally identifing records about you are maintained or if

you feel that your confidentiality has been compromised.
<br>     [comment: There should be every effort to make this administrator accessible to the public for questions and assurances, through email, phone, and personal contact.  This administrator should oversee the enforcement of library policies and keep the staff informed about these policies.]
<br>     [comment: There should also a brief disclosure of any state or local laws which oversee the confidentiality of library records or potentially overrule these polices, such as "sunshine laws."]
<p><a NAME="Links"></a>Links
<p><a href="http://www.ala.org/alaorg/policymanual/index.htm">The ALA's Policy Manual</a>
<br><a href="http://staffweb.library.vanderbilt.edu/ala_tf/report.htm">ALA Task Force Report on Privacy in the Online Environment</a>
<br><a href="http://www.kcoyle.net/privacyprimer.html">A Primer on Internet Privacy</a>
<br><a href="http://www.epic.org">The Electronic Privacy Information Center</a>
<br><a href="http://www.epic.org/reports/surfer-beware.html">Surfer Beware: Personal Privacy and the Internet</a>
<br><a href="http://www.junkbusters.com/cgi-bin/privacy">Junkbusters' Privacy Alert Page</a>
<br>
</body>
</html>

## APPENDIX 3: HTML SOURCE OF THE POLICY ANNOTATION

```
<!doctype html public "-//w3c//dtd html 4.0 transitional//en">
<html>
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
  <meta name="Generator" content="Microsoft Word 97">
  <meta name="Template" content="C:\PROGRAM FILES\MICROSOFT
OFFICE\OFFICE\html.dot">
  <meta name="GENERATOR" content="Mozilla/4.73 [en] (Win95; U) [Netscape]">
  <title>Online Privacy</title>
</head>
<body link="#0000FF" vlink="#800080">
 
<br>
<center>
<p><b> I. INTRODUCTION</b></center>

<p>     "Privacy" is a complex concept which has taken many
meanings over time.  In the past, privacy was understood to be a right to freedom from
observation and scrutiny in the personal life of an individual, so that privacy was felt to be
violated when a person's behavior was observed or otherwise made public in a way the person
found invasive and revealing.  Thus, an individual's life within the home and among
intimate acquaintances was understood to be protected to some extent from public
scrutiny.  But nowadays privacy is also thought to extend to information about a person,
especially information which related to behaviors or status that is not regarded as public. 
Thus financial records, health records, and the like are thought to be "private" in the sense that
such information is or should be under the control of the subject of this information.  It is
also commonly held that intimate, personal exchanges between an individual and various
professionals who are given personal information about the individual should be regarded as
confidential.  Because information seeking behaviors can potentially reveal private
information about the seeker, and because such information seeking often is assisted by
information professionals with an understanding of confidentiality, librarians and other
information professionals must recognize and protect the confidentiality of any records of
information seeking.  This paper will explore guidelines outlining the protection of
privacy and specific kinds of information that are generated in information seeking, in order
to develop a set of recommendations for the protection of the confidentiality of library users in
the online environment.
<br>
```

<br>
<br>

<p>II. GENERAL STATEMENTS OF POLICY AND THE SCOPE OF CONFIDENTIALITY/PRIVACY

<p>  Confidentiality and privacy are values which have been linked to both freedom of expression and the right to access information.  Librarians and information professionals in general recognize the importance of maintaining the trust of information seekers, and have undertaken a variety of efforts to insure that the confidence and privacy of their clients are protected.  As the internet and online environments have become increasingly important sources of information, they have also created a number of new contexts for privacy issues.  This paper will identify and explore the privacy issues which accompany online information seeking, and what steps librarians should take to address these issues.  To see that privacy is of importance to librarians, this paper will exhibit the policies of the ALA and some statements made by individual librarians and information professionals. Then, what privacy protection consists in will be identified, and finally these principles will be applied to specific issues in the online environment.
<br>
<br>
<br>

<p><a NAME="alapolicies"></a>Privacy and Confidentiality Policies of the ALA

<p><br>
<br>
<p>     The American Library Association's policy manual, section 54.16 "On Professional Ethics," offers a series of statements to guide ethical conduct for librarians, one of which is "We protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired, or transmitted."   This simple statement establishes that a core value of librarianship is the protection of every user's privacy.  The library user may not think very much about confidentiality when seeking information, but even so there is a duty to safeguard the user's privacy and confidentiality.  It is not this paper's purpose to explore the reasons and justifications supporting this duty on the part of libraries, although it is clear that in part this duty arises from the potential "chilling effect" a lack of protection might create among users.  The important point is that the ALA recognizes confidentiality and privacy as important rights of library users.
<p>     One area where the confidentiality of library users has been challenged is in circulation records.  Law enforcement agencies have requested this information, either to obtain the records of specific persons or as part of "fishing expeditions" to identify potential suspects.  The ALA has repeatedly fought such efforts and section 52.4 specifically addresses the confidentiality of library records.  In fact, 52.4 identifies a variety of records, beyond just circulation records, for protection:

<blockquote>Confidentiality extends to "information sought or received, and materials consulted, borrowed, or acquired," and includes database search records, reference interviews, circulation records, interlibrary loan records, and other personally identifiable uses of library materials, facilities, or services.</blockquote>

There are several important ideas in this sentence.  First, it states that confidentiality extends to any library record that could identify particular persons.  Second, it states that confidentiality extends to uses of library resources, suggesting that any use of resources which produces some type of record is to be protected, not just the uses of resources that are mediated by library personnel.  Thus, if a library user were to use a library computer to seek information, any electronic records of such use (and we shall see that there are several) should also be protected. <a NAME="aggregatenotpersonal"></a>Finally, it should be noted that there is an emphasis on personally identifiable information -- that is, information which can be linked with specific individuals.  Where records of information seeking are aggregated, abstracted, or otherwise collected into data that does not identify specific individuals in a way that can be traced back to individuals, this information can rightly be regarded as falling outside the concerns for and protections of privacy and confidentiality.

<p>     Section 52.4 of the policy manual ("Confidentiality of Patron Records") addresses the appropriate legal procedure for law enforcement to access such confidential records, but it also stresses the need for libraries to adopt formal policies recognizing the above standard of confidentiality.  In order to adequately address the confidentiality of records, libraries must address the specific kinds of records created in the online environment, which will be discussed below.

<p>     The ALA's <u>Intellectual Freedom Manual</u> reiterates this policy under the headings "Policy on the Confidentiality of Library Records" and "Policy concerning Confidentiality of Personally Identifiable Information about Library Users."  These two sections actually present portions of 52.4, and also provide some historical background on the events which led the ALA to issue the policy.  Apart from the historical accounts, the Intellectual Freedom Manual does not alter or expand the scope or range of records identified in the ALA Policy Manual.

<p>     Apart from the general policies of the ALA, there is another document outlining the impact of computer technology on the privacy and confidentiality of patron records.  This document is the <u>Report of the Task Force on Privacy and Confidentiality in the Electronic Environment</u>, issued July 7, 2000.  The report includes a number of important findings.

<p>     First, there is the issue of computer system security.  The Task Force found that insofar as some patron records exist as electronic records on library computers, the library must recognize a duty to protect these records from hackers or other unauthorized access.  Specific security measures fall outside the scope of my work but some account of the library's commitment to resist unauthorized access to computer records should be disclosed to users.

<p>     Second, the Task Force found that a variety of privacy issues are raised by in-library internet access.  Every internet transaction is potentially a source of information gathering by web sites and servers outside the library.  Because many patrons may use a computer over the course of a day, there needs to be some awareness of the potential for subsequent users to see traces of the web sites visited by previous users. 

There is also the potential for other patrons to see the screen of a computer being used by a patron, thereby compromising the user's privacy.
<p>     Third, the Task Force recognizes potential problems raised by the use of electronic resources provided by a remote vendor.  These vendors may collect data about resource access and even about the patrons accessing them without guaranteeing the same protections of a library.  Thus, it is important for libraries to establish privacy protection measures in their licenses and contracts with vendors or else to give patrons notice that personal information may be collected by the vendor.
<p>     Apart from these significant findings, the Task Force issued a number of recommendations.  Beyond revising ALA policies concerning confidentiality of records to address the findings mentioned above, the Task Force calls on the ALA to "urge that all libraries adopt a privacy statement on web pages and post privacy policies in the library which cover the issues of privacy in internet use as accessed through the library's services."
<br>
<br>
<br>

<p>The FTC's Online Privacy Reports
<p><br>
<br>
<p>     The Federal Trade Commission's reports on developing a theoretical framework for fair information practices online stand out as particularly explicit and well-informed documents.  The focus of the FTC's two reports to Congress has been on commercial use of the internet, but the general discussion of fair information practices here is particularly useful.  The FTC reviewed some twenty-five years worth of government documents and studies concerning the collection of personal information and identified five core principles of privacy which were common to all.  These principles were identified in 1998 as Notice/Awareness, Choice/Consent, Access/Participation, Integrity/Security, and Enforcement/Redress.  Before discussing the particulars of each principle, it should be noted that the FTC's 2000 report to Congress pares down the list Notice, Choice, Access, and Security -- Enforcement/Redress is dropped from the core -- presumably because the 2000 report is intended to guide enforceable legislation, and not because the FTC no longer views the enforcement of privacy protection as important.
<p><a NAME="NOTICE"></a>     The principle of notice is regarded as the most fundamental of the principles.  The idea is that the individual must have an awareness of the practices of information collecting agencies before he can decide to disclose personal information.  In addition, the individual would need to know that information is collected before he could attempt to access, evaluate, or correct the information, and also before he could make any effort to redress illicit or unauthorized use of the information.  The FTC concludes that the principle of notice requires that personal information collecting entities should be required to disclose at least some of the following in order to provide meaningful notice: (1) the identity of the entity collecting data; (2) the uses to which data is put; (3) the identity of any potential recipients of the data; (4) the nature of the data collected and means of collection; (5) whether the individual user or consumer is required to furnish personal information, and the consequences of refusal to do so; and (6) the measures

taken by the collecting entity to protect the data, ensure its accuracy, and uphold its confidentiality. The FTC's requirements are not quite stringent enough for libraries, in so far as libraries have a positive commitment to protecting confidentiality, and so all of these items should be disclosed in a library privacy policy.  The FTC does however raise an important point about the posting of policies: they should be readily assessable both from the home page of a web site and from any pages where personal information is collected.

<p><a NAME="CHOICE"></a>     The principle of choice requires that the individual whose personal information is to be collected be given the option to furnish information or not.  It also requires that the individual be given a choice about how the information will be used, and especially about how the information may be put to use beyond the initial transaction that requires the disclosure of personal information. Thus, if the owner of a web site intends to collect personal information, there must be consent from the user or consumer both for the initial collection of information and for any uses the information might be put to: web site management and development, sale to advertisers or direct marketing firms, sharing with other companies for market research, or any other use.  This principle can be implemented in one of two ways -- either allowing the user to opt-in or to opt-out. <a NAME="OPTIN_OUT"></a>"Opt-in" means the individual must actively consent to the collection and use of information.  "Opt-out" means that the user must actively deny consent for information collection and use.  It is hardly surprising that most collectors of information prefer the latter, as it allows them to assume consent from any individual, and requires special effort on the individual's part, rather than on the information collecting entity's part, to protect privacy.  Libraries, however, would probably show greater fidelity to the user's confidentiality by using "opt-in" protocols wherever possible, as this provides an additional opportunity to both disclose policy and assure the user of the library's commitment to privacy.

<p><a NAME="ACCESS"></a>     The principle of access is comprised of two rights on the part of the individual about whom information is collected: the right to review any information and the right to correct or contest the information. This principle can only be upheld effectively if there is a relatively easy and inexpensive procedure for reviewing and correcting the data.  The more this information is limited in scope and content, the easier it will be to ensure that the information is accurate.

<p><a NAME="SECURITY"></a>     The principle of security covers both the integrity or quality of information and the protection of information from loss and the unauthorized disclosure, use, and access of the information.    The integrity of information is protected in part by allowing the individual to access and correct information, but there is also a responsibility on the part of the collecting entity to use prudence in selecting sources of information.  Finally, the occasional destruction of obsolete or old information is necessary to protect the quality and integrity of information.  Protecting the security of information calls for both technological and administrative measures. The technological measures involved might include encryption, pass-word protection, firewalls, and storage on secure servers.  The administrative measures should focus on setting limits on who can access the information and to what purposes the information can be used.  Obviously, it is imperative that the information never be used in manners or for purposes not authorized by the individual.  Libraries should have policies in place which limit access to patron records of any kind, and also have technological protections sufficient to prevent unauthorized access to records.

<p><a NAME="REDRESS"></a>     The last principle is the principle of enforcement.  The FTC concludes that "the core principles of privacy can only be effective if there is a mechanism in place to enforce them." The FTC outlines several mechanisms of enforcement: self-regulation, private remedies, and government enforcement.  Self-regulation, the FTC contends, can be carried out by requiring members of professional or industrial organizations to accept and comply with a privacy policy.  There is a clear application of this principle for libraries: the adoption of privacy policies by organizations such as the ACRL, the ALA, and the SLA.  These policies should be fairly explicit, and there would also be a need for developing a system to verify compliance.  Because it is unlikely that libraries would be motivated to violate privacy policies (unlike, say, private firms which might realize a profit from selling customer information), it may be enough to simply require libraries to adopt formal privacy policies and post them on their web sites.  Tort law provides for a certain amount of private remedies.  But it is unclear, at this point, what steps legislatures may take to pass laws concerning privacy protection.
<p>     Taken together, the five principles of privacy protection identified by the FTC provide a strong foundation for specifying the principles libraries should uphold with respect to personal information.  But it is also helpful to notice to some specific statements of principles offered by librarians and other information professionals concerning privacy and confidentiality.  This is important to do because the FTC's principles constitute a minimal set of guidelines to restrict the activities of commercial entities.  But librarians are committed to doing more than merely setting limits on their use of personal information -- they must take positive steps to ensure that confidentiality is always protected.
<br>
<br>
<br>

<p>Guidelines of the Organization for Economic Cooperation and Development
<p><br>
<br>
<p>     Although it is impossible and unnecessary to discuss all of the many statements of principles of privacy protection created by organizations and individuals, it is enlightening to review the guidelines issued in 1980 by the OECD.  These guidelines consist of eight principles which have influenced many later lists of principles. It is possible to "map" these eight principles onto the five FTC principles to emphasize the wide area covered by the FTC's principles and to show that those five principles are sufficient to cover many privacy concerns.  The principles of the OECD consist of a collection limitation principle; a data quality principle; a purpose specification principle; a use limitation principle; a security safeguards principle; an openness principle; an individual participation principle; and an accountability principle.
<p>     The collection limitation principle calls for limits to the personal information collected. It also requires that the subject be informed of the collection, and when possible that consent be obtained from the subject.  This principle thus falls under the FTC's notice/awareness principle.  The data quality principle requires that

collected data be kept accurate and up-to-date, and so it falls under the access/participation principle.  The purpose specification principle requires that information gathered must have the purpose for its collection and use specified at the time of gathering, and the information must be put to no other use.  This falls under the security/integrity principle.  The use limitation principle requires that consent be obtained from the data subject should the data be used for other purposes than those agreed to, or should the data be disclosed to others.  This falls under the choice/consent principle and also the security/integrity principle.  The openness principle requires that information collectors be open about their policies, practices, and developments with respect to personal information.  It also requires that the collector respond to inquiries from data subjects and identify themselves.  This falls under access/participation and notice/awareness. The security safeguards principle requires that data be protected from loss and unauthorized access, use, or disclosure.  This falls under the security/integrity principle of the FTC.  The individual participation principle provides the subjects of data collection with the right to know whether or not data has been collected about them; to access the data about them; and to challenge and correct data about them.  This falls under the access/participation principle.  The accountability principle requires that data controllers be held accountable to these requirements, through legal or other means, and falls under the enforcement/redress principle.
<p>     It is clear, then, that the OECD principles, which were widely adopted before the emergence of the internet as it is now known, are exhaustively covered by the FTC principles.
<br>
<br>
<br>

<p>Personal Statements of Privacy Principles
<p><br>
<br>
<p>     Beyond the principles outlined by various organizations, several statements of privacy principles made by individual librarians and information professionals are worth noting.  Although these statements do not attempt to cover every aspect of privacy and confidentiality, they do focus attention on issues which are of particular importance to librarianship.
<br>
<br>
<br>

<p>Featherman's Ten Commandments
<p><br>
<br>
<br>
<br>
<br>
<p>     John Featherman of ASIS publishes a newsletter on privacy for consumers, and at the 1997 ASIS mid-year meeting presented a set of rules for individuals to help ensure that their privacy was protected.  Although many of these focus on strategies

for individuals, there are several rules that are applicable to institutions as well.  His rules -- he calls them the "Ten Commandments of Privacy" -- are:
<ol>1. Thou shalt keep sensitive information private.
<br>2. Thou shalt pay in cash whenever possible.
<br>3. Thou shalt guard thy social security number and other identification numbers with thy life.
<br>4. Thou shalt use a paper shredder in thy daily life.
<br>5. Thou shalt use a post office box or, better yet, a mail drop.
<br>6. Thou shalt inspect thy credit, medical, and other personal information files often.
<br>7. Thou shalt be circumspect in thy computer affairs.
<br>8. Thou shalt be extremely discreet when communicating.
<br>9. Thou shalt be diligent when choosing passwords and shalt change them often.
<br>10. Thou shalt make a lifetime commitment to protecting you privacy.</ol>

Each of these commandments is clearly directed at individuals rather than institutions, but there are several points raised here that are worth considering.  The third commandment should remind us that Social Security Numbers, driver license numbers, and other personally identifying numbers must be guarded with special attention.  Many libraries never request this kind of information, but some may use them in patron records.  For example, the Kent State University library uses student ID cards as library cards, and the student number printed and encoded on each card is usually the individual's Social Security Number.  The third commandment should call attention to the sensitivity of these student ID numbers.
<p>     The fourth commandment is important at an institutional level because paper records are often produced in libraries, as in the ILL and book request process.  Computer print-outs of web pages produced inside a library may also be potentially identifiable, because the most popular browsers include the time and date as headers or footers in print-outs, and those libraries that require registration to use computer terminals make it possible for a persistent and industrious snoop to identify the patron who printed a particular page.  For this reason, abandoned print-outs should be destroyed and not merely thrown out.
<p>     Commandment six calls on individuals to exercise their right to access information collected about themselves, as outlined in the FTC's principle of access.  Commandment nine calls attention to the need for librarians to periodically change the passwords that provide access to personal information -- and the principle of security holds that we ought to restrict access to personal information with passwords.
<p>     Finally, commandment seven focuses on computers as potential privacy hazards.  Featherman explains that encryption should be used whenever personal information is transmitted, and we should add that libraries would do well to use encryption whenever they transmit information that may identify patrons, such as ILL requests.
<br> 
<br> 
<br>

<p>Gorman's "New Libraries, Old Values"

<p><br>

<br>
<p>     Michael Gorman lists privacy as one of eight values of that underpin librarianship.  He writes:
<blockquote>The confidentiality of library records is not the most sensational weapon in the fight for privacy, but it is important, both on practical and moral grounds.  In practical terms, a lot of the relationship between a library and its patrons is based on trust, and, in a free society, a library user should be secure in trusting us not to reveal what is being read and by whom.  On moral grounds, we must start with the premise that everyone is entitled to freedom of access, freedom to read texts and view images, and freedom of thought and expression.  None of these freedoms can survive in an atmosphere in which library use is monitored and individual reading patters are made known to anyone without permission.</blockquote>
Gorman quite rightly focuses attention on the role of trust in library patronage, and speaks in technologically neutral language about the right to read text and view images with confidence that one's information seeking will not be revealed to others.  In the context of online information seeking, this means that a user should never be identified with web pages or computer files accessed.
<br>
<br>
<br>

<p>Coyle's Elements of a Privacy Policy
<p><br>
<br>
<p>     Karen Coyle, a privacy advocate and member of the ALA's Task Force on Privacy and Confidentiality in the Electronic Environment,  has outlined essential elements for a library web site privacy policy. Her elements fall under four main categories: Background Information, Your Library's System, Policy Elements, and Links.
<p>     Background information should include both legal and policy statements.  Coyle recommends providing information about state and local laws concerning public records and library records.  This seems reasonable, as users may find it reassuring to know what protections exist.  But they will also find it useful to know what limits there may be in the protection of confidentiality where laws diverge from library values.  The policy statements would include both the ALA's policy on confidentiality and local policies -- that is, the policy of the institution to which the library belongs and the specific policies of the library.
<p>   The section on your library's system would outline the specific issues raised by the network, web page design, and connected but remote systems.  These system elements would include a variety of matters.  The logging system the library uses on its web server, which tracks http requests and browsing, should be explained.  Likewise, any personalization features on the system should be explained so that users would know what information about them is stored in the system, such as their log-in history, e-mail features, and document delivery.  Finally, the user should be made aware that personally identifying information may be requested by remote databases accessed through the library system or furnished to interlibrary loan partners.
<p>      The policy elements would contain details

about the access, integrity, and security of data stored on the library system.  The protections in place, the length of retention, and who can access the data should all be disclosed.
<p>      Lastly, the policy should include links to further clarify and explain privacy issues.  Coyle specifically recommends that the online ALA policy manual should be linked, as well as the ALA task force report on online privacy.  This list of links should not be considered exhaustive, but note that protecting user privacy and providing privacy literacy training can be distinguished.  It is beyond the scope of my project to create a privacy literacy curriculum.
<br>
<br>
<br>
<br>
<br>

<p>III. SPECIFIC PRIVACY ISSUES OF THE ONLINE ENVIRONMENT
<br>
<br>
<p>Cookies
<p><br>
<p><a NAME="COOKIEdf"></a>     A "cookie" is small text file -- 4k or smaller -- which records information on the hard drive of a computer requesting a web page.  The content of a cookie generally consists of three main elements: the address of the server or web site which sent the cookie, an identification code for the specific computer, and a set of codes recording information about the user browsing habits and/or preferences.  The contents of the cookie are encoded so that the recipient is unlikely to be able to interpret what they actually contain.  Here is an example cookie from my own computer, the filename of which is "anyuser@fedex[1].txt":
<blockquote>FedEx_accrue
<br>12000206511118www4744177
<br>fedex.com/
<br>0
<br>2758148096
<br>31550824
<br>1850881248
<br>29380962
<br>*</blockquote>
     The above contents are meaningless to me, although I can recognize the server domain "fedex.com."  Presumably the codes are used by FedEx to personalize their web pages when I view them, or could be used to keep track of goods or services I requested.  But in order to protect the privacy of web surfers, there ought to be both notice that this information is collected, and some way to insure that the information is accurate, if this information is in any way identifiable with the surfer.
<p>     It should be noted that some cookies will use potentially identifying user names in the filename.  For example, many cookies on my computer have filenames like "mike@nytimes.com."  The cookie used my Windows logon name, which

is "mike," as my userid.  Because many computer networks require that users log on with some name, this method of naming cookie files presents a minor privacy threat.  Many web servers will maintain databases of users and their browsing habits, and these userids are sometimes based on the user's actual name ("mmonaco" for example).  Any such cookies certainly seem to constitute a personally-identifying record, even if the identification is would require a bit of sleuth work to connect it to the actual user.

<p><a NAME="COOKIEtypes"></a>     There are two kinds of cookies: session and persistent.  A session cookie is deleted by the browser as soon as the cookie sending server is exited.  A persistent cookie will remain on the hard drive for a specified period of time.  The only difference between session and persistent cookies is that persistent cookies, when set by the server, have an additional variable in the command which sends the cookie: "expires=date." Session cookies are somewhat less threatening to privacy than persistent cookies because they are deleted after every session, thus limiting the amount of information a particular web site collects about the user. However, because many servers will keep their own databases of user information, there is still a need to disclose the use session cookies to users in order to protect their privacy.

<p>     Most commercial web sites and many non-commercial web sites use cookies.  This is because there are many legitimate uses for cookies.  Cookies which track the browsing habits of users provide information useful for improving and designing web sites.  Cookies can also add several features to web sites, such as custom formatting of web sites (some users may request specific colors, fonts, or font sizes for easier reading), alerts about new material on the site (a persistent cookie would record the last visit of a user and allow the server to provide a list on material since that last visit), and "shopping baskets" to keep a record of items requested, online forms completed, and the like.  Disabling cookies (an option provided by browsers which use them) can therefore create a loss of functionality, and some web sites would be unusable without cookies.

<p>     Cookies present several privacy issues to users of online resources.  At the most basic level, cookies may be considered a privacy threat because they can operate as "spies,"  in the sense that the senders of cookies may not reveal themselves while at the same time collecting information about the user.  An informal study by Karen Coyle of cookies collected on her hard drive revealed that about one third of the cookies were untraceable, owing to the fact that the web addresses in the cookies could not be reached. The cookie senders thus remain anonymous, which in itself raises questions about the intentions of the cookie senders.

<p>     It has been noted that cookies, by themselves, do not actually identify the user who receives them.  Cookies simply collect clickstream data, indicating what a user views, how long he or she is connected to the sites, and which advertisements were displayed by the site.  By itself, this data does constitute useful information for the web site designer and marketer.  The real privacy threat emerges when cookie technology is combined with another technology: online forms.  Users may be asked for identifying personal information including addresses, names, phone numbers, and the like when they register to use various web sites, services, or files.  If this information should be matched with clickstream data collected from cookies, the user has suddenly revealed much more than the information voluntarily released in the form.  A profile of the users' interests can be determined by studying the clickstream data from cookies which reveal where, when, and for how long the user has surfed web sites.  The appeal of such information for marketers and

advertisers is obvious.  But the users' privacy has been compromised if the information collected by cookies is collected covertly, without disclosure.

<p>     It is believed that most cookies can be accessed only by the site which created them.  However, a bug in Microsoft's Internet Explorer was discovered in May of 2000, called the "Open Cookie Jar."  This bug allowed any server to read any cookie, provided that the name of the site which created the cookie was known. Microsoft has, since then, corrected the problem, and even offers cookie management features in the newest version of Internet Explorer.  Even so, this episode calls attention to the potential danger that cookies be intercepted or viewed illicitly.  A similar bug was found in one version of Netscape Navigator.  Bugs such as these clearly violate the principle of security, as they allow unauthorized disclosure and use of the information in the cookie.

<p>     It is possible for the server which sends cookies to specify that cookies will be sent only over secure communication channels; the set-cookie http response header would need to include the command "secure."  Presently, only HTTPS servers (http servers on SSL) are considered "secure" by the Netscape browser.

<p>     One further privacy threat posed by cookies lies in the fact that some banner advertisers send cookies as well.  Because the server providing banner ads sets the cookie rather than the server providing the actual page visited, these banner ad companies are able to track a user's browsing across many different sites and servers.  The purpose of these cookies is to aid in targeting specific ads at users.  However, because these companies store information about the recipients of the banner ads and their browsing habits, the user's privacy is seriously threatened, especially if this information is combined with forms which ask for personal information.  It is one thing if Doubleclick.com tags my computer with a cookie that records my browsing habits as "mike@doubleclick.com" and quite another if they should be able to connect my userid with my name and address because I filled in a form to register to use the online New York Times site.

<p><a NAME="COOKIESuse"></a>     Libraries should consider cookies carefully.  Computers provided for patron use will accrue a large collection of cookies.  The risk to patron privacy lies in the danger that when a patron fills in an online form, their personal information may be erroneously linked to the browsing habits of other users.  But personally identifiable information must be accurate, according to the principle of integrity discussed above.  On the other hand, patrons may be accurately profiled and their privacy compromised if they are unaware that the library's computers accept and use cookies.  There is also the issue of library web sites using cookies.  Better access may be provided to users by allowing them to specify larger fonts or other formatting issues.  And librarians would be able to improve their web sites by reviewing usage.  But these benefits must be balanced against the danger that cookies from library web sites could be intercepted or accessed illicitly.  The danger of interception may be reduced by always specifying that cookies be sent only over secure communication lines.  The danger that others may view cookies, for example by exploiting the Open Cookie Jar bug, is probably best met by using only session cookies.

<br>
<br>
<br>

<p>Clickstream and Log Files

<p><br>
<br>
<p>     Apart from cookies, there are other traces of internet use recorded by web servers and private firms.  These other records are loosely grouped under the terms "clickstream," "mouse droppings," and "data trail."  This discussion will use the term "clickstream" as it is the most widespread term.
<p><a NAME="HEADERS"></a>One source of clickstream data is found in the headers sent by web browsers to servers when a user accesses a web page.  These headers reveal the internet protocol address ("IP address") of the user, the time of day when the page is accessed, the pages and images downloaded, the page from which the user is accessing the new page ("Referrer"), and data entered by the user into forms.  Cookies are sometimes considered to be part of this clickstream data as well, but cookies will not be explored further here.
<p><a NAME="LOGFILES"></a>In order to give a picture of what one might look like, a sample log file from Karen Coyle's web pages is reproduced here:
<blockquote>ppp156210.asahi-net.or.jp -- [29/Mar/1998:05:45:00 -0800] "GET /~kec/best.html HTTP/1.0" 200
<br>DIAL8-ASYNC38.DIAL.NET.NYU.EDU -- [29/Mar/1998:12:23:11 -0800] "GET /~kec HTTP/1.0" 302
<br>gwc-mel.afwgc.af.mil -- [29/Mar/1998:14:24:09 -0800] "GET /~kec/ HTTP/1.0" 304</blockquote>
As can be seen, an IP address is recorded, and then the date and time that pages are accessed, and finally the path of the page accessed.  Each of three entries follows the format "IP address -- date/time -- command/request."
<p>     All of this data is provided by the web browser used to access the pages.  But the individual using the browser is likely to be unaware that this information was provided to the server visited.  The consumer advocacy web site "Junkbusters.com" has a page which allows users to see for themselves what their browser tells every server they visit.  When visiting the Junkbusters page with a computer in a KSU philosophy department, the following output is produced:
<blockquote>The ``HTTP Referrer" tells them what led you to the request. In this case it was http://www.junkbusters.com/junkdata.html.
<br>Some advertisers choose which ad to show you based on two variables that tell them which computers you are coming through: the ``Remote Address" (in your case <b>131.123.68.123)</b> and the ``Remote host:" (in your case 131.123.68.123). These variables may give indications of where you live or work.
<br>The ``User Agent" variable, in your case Mozilla/4.73 [en] (Win95; U), indicates what software and hardware you are using. This interests companies that sell these goods. (Mozilla is Netscape's name for its browser.)
<br>Some browsers actually hand over your email address or other indications of your identity. In your case it appears the ``HTTP From" variable was <b>not provided</b>, and the ``REMOTE_USER" variable was not provided.</blockquote>

This URL basically displays the headers sent by the browser when a user visits the address -- they are displayed above in bold face.  The text on the page also explains the specific

variables used by http to record this information.  Junkbusters encourages people interested in privacy to link to this page in order to promote privacy literacy (and to promote their privacy software).  Libraries would do well to make this page available to their users.  But perhaps the most interesting thing to notice here is that there are header variables not recorded in Karen Coyle's sample log entries but which other servers might choose to log.

<p>     Another source of clickstream data is in the browser "add-ons" known as Java, JavaScript, and VB Script.  These are all technologies which enhance the web pages with graphics, animation, and the like.  These will generally provide information about the browser being used, screen resolution, and other data of concern mainly for formatting purposes.

<p>     The information sent back and forth in headers is of little concern on its own, but a threat to privacy emerges when this information is stored for later use and analysis in transaction log files.

<p><a NAME="IPADDRESS"></a>One thing to notice is that the clickstream data stored in log files would identify computers but not individuals.  However, an IP address can reveal a certain amount of information about an individual.  IP addresses are assigned by the internet service provider (ISP), and often disclose the region of the country or even the city the computer is in. A quick review of the IP addresses in Karen Coyle's log file suggests that the three transactions logged were requested by users in Japan,  New York University, and a military base. IP addresses that consist entirely of numbers are in principle traceable as well. The IPs also reveal the identity of the ISP, which in itself reveals little about the individual.  The real danger emerges when clickstream data is merged with personally identifying information.  If your identity can be connected with your clickstream, others can know what web sites you visit, which links you follow, what files you download or view, and so on.

<p>     Perhaps the most egregious and high profile case of compromising privacy in this manner is the recent controversy surrounding DoubleClick.com.  DoubleClick.com acquired Abacus Direct, a company which maintains databases of consumer information compiled by direct marketers.  The kind of information Abacus Direct maintains involves consumer habits and demographics.  DoubleClick.com provides banner ads to thousands of web sites, and as described above, these banner ads send cookies to web surfers who see their ads.  These cookies identify the user to DoubleClick.com, and the browser headers transmitted when surfers click on the banners are compiled in DoubleClick's log files.  When the Abacus Direct databases are combined with DoubleClick.com's log files, the resulting database identifies individuals with demographic information, names and addresses, web surfing habits, and consumer information.  All of this information has obvious use for advertisers who want to target ads effectively.  But this information also constitutes a severe compromise of the surfer's privacy, not least because the individual is unlikely to be aware of the clickstream information collected.

<p>     A persistent problem with online privacy is the fact that so many users are unaware of the information they are providing to web sites and servers.  Online forms represent the only conscious, voluntary release of information for most web users.  The average web user simply can't know what cookies to accept or reject, and almost certainly does not realize a trail is created whenever he or she clicks on a link. 

Again, the principle of notice requires that the surfer be notified that header information is logged.
<p>     DoubleClick.com, in the face of criticism from consumer advocates and privacy advocates, has promised to await the development of internet privacy rules before merging the log files and databases.  But it is evidently just a matter of time before they or other companies will begin to create detailed profiles of web users.
<br>
<br>
<br>

<p>Email
<p><br>
<br>
<br>
<p>     Because some library patrons use email to ask reference questions, and because some libraries offer internet access which is used by patrons to access email accounts, email policies are often adopted by libraries or their parent institutions.  These policies should include a privacy element for several reasons.
<p>     First, the Electronic Communications Privacy Act, passed by Congress in 1986, guarantees certain protections from the interception and disclosure of email messages.  Just as the confidentiality polices of the ALA call on libraries to protect patron records and disclose them to law enforcement or other government agencies only in cases were good reason and proper legal documentation are provided, the email messages of library patrons must be protected.
<p>     But many institutions place limits on the appropriate use of email accounts, and libraries may do the same.  The crucial point is that any policies must be disclosed to users.  So if the library monitors email messages created on library computers or sent to reference personnel, these practices must be disclosed in the privacy policy statement.
<p>     Lastly, because emailed reference questions are created by patrons in the information seeking process, they constitute records that ALA policies require libraries to protect as confidential.
<p>     In order to adequately protect user confidentiality, then, the online privacy polices of libraries must disclose any monitoring of email.  Presumably, any such monitoring would be undertaken only in order to assure that library resources are not abused or as part of quality assurance in the case of reference questions.  The library should not keep unnecessary records of messages sent and received, and protect those it records as confidential.
<br>
<br>
<br>

<p>Interlibrary Loans
<p><br>
<br>

<p>     Requesting and providing materials for interlibrary loans (ILL) is another library activity which generates records.  These records are likely to identify the patron requesting materials.  Christopher Nolan, writing in the Journal of Academic Librarianship, notes that the paper ALA-sanctioned ILL request form and the online work forms used by OCLC, RLN, and WLN all use information about the requesting patron, and supply this information to the institution making the loan.<sup> </sup>This creates a privacy issue for patrons because both the institution making the request on their behalf, and the lending institution, have records of the patron's information seeking.
<p>     There is obviously good reason to keep track of circulation, but Nolan points out that several things may happen to the ILL request.  Because the patron is identified in ILL requests, traces of the loan will remain even if the library conscientiously destroys old circulation records.  First, the CONTU guidelines attached to copyright laws require that records of all photocopies of materials obtained through ILL be maintained for three years.  Commonly, these records are maintained by keeping a copy of the ILL form which generated the photocopy. Because the patron is identified on this form, there is now a record of materials borrowed which is identifiable with the patron.
<p>     Second, the ILL request forms are supplied to the lending institution.  Again because the forms contain the patron's name and possibly other information, the lending institution now has a record of the patron's information seeking.
<p>     Privacy policies must address these records.  An ideal policy would require that CONTU guidelines be met by keeping a non-identifying record of photocopies made and would require that patron information not be sent to the lending institution.  Because some ILL forms are online, it would be a good idea to disclose the library's policies about ILL requests in the online privacy policy.
<br>
<br>
<br>

<p>Searches
<p><br>
<br>
<p>     Another privacy issue of the online environment is internet searching.  There are two ways that searching the internet can threaten a library patron's privacy.  First, many search engines and directory sites will record the actual search terms entered by the user in the URL of search results.  For example, doing a simple search of the term "clickstream" at Yahoo and Excite lead my browser to the following addresses:
<blockquote>http://search.yahoo.com/bin/search?p=clickstream
<br>http://search.excite.com/search.gw?search=clickstream</blockquote>
If I were to follow the links provided on this page, my browser's header would disclose this URL as the "HTTP Referrer" variable.  Thus, my search strategy is potentially recorded in a log file, and it is connected with my IP address.  If a library web site provides links to search engines which encode search terms in the results URL, this should be mentioned in the privacy policy.
<p>     Second, many search engines use cookies to keep track of users and their queries.  This is not a privacy problem when the search engine keeps only

"aggregate" information -- that is, statistics about user trends and surfing habits.  But persistent cookies used by search engines to track a particular user's queries and searches do threaten privacy.  As discussed above under the topic of cookies, there is always a risk that cookie information will be linked with the subject's identity when the subject completes a form or otherwise releases information online.  And as in most online privacy issues, users may be unaware that information is being disclosed.

<p>     Apart from internet searches, libraries need to consider the possibility that OPACs can retain digital traces of the searches conducted by users.  Users occasionally leave a terminal with their search results and strategy visible on the monitor.  In so far as the patron has chosen to leave the terminal without returning to the menu screen or otherwise concealing his or her search, there is little the library can do to protect the confidentiality of searches from this kind of disclosure.  But libraries will also have to consider the possibility that the OPAC retains records of previous searches in its memory, disk cache, or log files.  Whenever this is the case, the policy of the library must be to ensure that these records are secure from unauthorized access and destroyed as soon as they serve no role in the library's functioning.  At the very least, records of searches maintained for statistical analysis should have no personally identifying information.

<p>     In most cases, OPACs do not require logging in or identifying the user.  But there are exceptions, most notably in systems which allow users to view their own circulation records.  Because this confidential information is accessible, it is the library's responsibility and duty to take adequate measures for protecting the information.  Circulation records should be protected by passwords.  They should also be transmitted to authorized accessers in an encrypted format to protect against interception.

<br>
<br>
<br>
<br>
<br>
<br>
<br>
<br>

<p>IV. CONCLUSIONS

<p>     The protection of privacy can be assured by meeting the five main principles outline by the FTC: notice/awareness, choice/consent, access/participation, integrity/security, and redress/enforcement.  Librarians must strive to uphold each of these principles in their handling of confidential information.

<p>     The information generated by online library use includes cookies, clickstream data (including the header data created by browsers and log files maintained by servers), ILL request forms, and email reference questions.  All of this information is potentially personally identifiable and thus should be protected by strict standards of privacy as outlined in the five principles.  When data is abstracted from log files into statistical data or otherwise severed from personally identifying marks, the data can be considered outside the umbrella of privacy protection.

<p>     Insofar as library web sites and online systems collect personally identifying information or provide it to outside agencies such as online search engines or databases, there must be a clear disclosure of this practice and adequate measures taken to protect patron confidentiality.

<p>    Because libraries also provide access to databases, search engines, and web pages which belong to other agencies, it is also important that libraries give notice to the patron that the privacy policies of the library are not shared by these outside agencies. The patrons should be reminded that they cannot assume resources accessed through the library are covered by the same guarantees of privacy that library makes about use of resources belonging to it.

<p>    All disclosures of policies should be made in a privacy policy statement that can be accessed easily by online patrons, and an ideal policy will consider all of the issues outlined in this paper.

<br>

<br>

<br>

</body>

</html>